ILLINOIS

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

PRODUCTION NOTE

University of Illinois at
Urbana-Champaign Library
Large-scale Digitization Project, 2007.

Library Trends

Analyses of Bibliographies

H. R. SIMON
Issue Editor

July, 1973

Library Trends

A Publication of the University of Illinois Graduate School of Library Science

LIBRARY TRENDS, a quarterly journal of librarianship, provides a medium for evaluative recapitulation of current thought and practice, searching for those ideas and procedures which hold the greatest potentialities for the future.

Each issue is concerned with one aspect of librarianship. Each is planned with the assistance of an invited advisory editor. All articles are by invitation. Suggestions for future issues are welcomed and should be sent to the Managing Editor.

Published four times a year, in July, October, January, and April. Office of Publication: University of Illinois Graduate School of Library Science, Urbana, Illinois 61801. Entered as second-class matter June 25, 1952, at the Post Office at Urbana, Illinois, under the act of August 24, 1912.

Subscription price is $8.00 a year. Individual issues are priced at $2.50. All foreign subscriptions and orders should be accompanied by payment. Address orders to Subscription Department, University of Illinois Press, Urbana, Illinois 61801. Editorial correspondence should be sent to the Publications Office—*Library Trends*, 215 Armory Building, University of Illinois Graduate School of Library Science, Champaign, Illinois 61820. Indexed in *Current Contents, Current Index to Journals in Education, Library and Information Science Abstracts, Library Literature, PAIS,* and *Science Citation Index.*

US ISSN 0024-2594

PRINTED IN THE U.S.A.

Library Trends

VOLUME 22 • NUMBER 1

JULY, 1973

Analyses of Bibliographies

H. R. SIMON
Issue Editor

CONTRIBUTORS TO THIS ISSUE

This Page Intentionally Left Blank

Introduction:

Why Analyze Bibliographies?

H. R. SIMON

IT IS becoming increasingly apparent that for the continued development of information studies it is above all necessary to investigate the paths along which scholarly information is moving. Since the scholarly exertions of antiquity, the flow of information has possessed vast reservoirs in the form of bibliographies. These represent a concentration of information such as may excellently serve, from the viewpoint of the statistical method, as representative samples for the investigation of the information flow as a whole. Bibliographies enable us to develop model formulations which may ultimately give rise to a comprehensive theory of information.

Bibliographies are worth investigating under carefully and methodically defined conditions. Two types of investigations could prove particularly helpful:

1. Investigations from the perspective of general information theory—the results of these inquiries would be of value to information specialists of all kinds. Classed in this area could be Martyn's article in this issue which gives an evaluation of the general situation regarding secondary literature; and Brookes's article which provides a detailed exposition of the methods of numerical analysis of bibliographies. An article was planned on the use of computer-stored bibliographies as sources for future developments in information and library science. Unfortunately this article did not materialize; however, an earlier *Library Trends* article by Malin describing the *Science Citation Index*®[1] might be helpful to the reader interested in this area.
2. Investigations from specialized viewpoints—the results of these analyses should reveal the current situation as well as trends in individual

H. R. Simon is Archivist and Lecturer for the Library, Zoological Institute of the University of Heidelberg, Germany.

areas of research. They are important for scholars in a great variety of specialized areas as well as for administrators in scholarship. In this issue the articles by Thompson on the humanities and Bottle on scientific bibliographies represent investigations from specialized viewpoints.

These two approaches to analyzing bibliographies, which naturally represent only crude classifications, must be kept in view when deciding to set up projects for the analysis of bibliographies. Lancaster has listed criteria important for the information scholar.[2] In addition to the well-known and frequently discussed factors of coverage, recall and precision, Lancaster distinguishes that of novelty. Novelty is a measure of the degree to which a bibliography brings to the user's attention references to articles that are new to him (i.e., articles that he was not aware of before seeing the bibliography). A novelty ratio[3] may be expressed in the form:

$$\frac{\text{Number of articles that are relevant to the user and new to him}}{\text{Number of articles that are relevant to the user}}$$

With the two criteria of novelty and costs, it may exactly be determined, above and beyond the general informational process, how a bibliography is rated from the viewpoint of the user. The fact, therefore, that there are two quite decisively important questions to be considered in the practice of information transmission may easily be shown by means of appropriate investigations.[4]

All the mentioned forms of analysis of bibliographies primarily concern information transmission. The conveying of information, however, takes place within one specialized context, within one certain field of research.[5] In view of this, bibliographical analysis is of importance to the scholar himself, although it can offer a valuable stimulus to the information specialist as well. To be more specific, we can ask the following questions, for example, in order to determine the information flow in a specialized area by means of analysis of its bibliographies:

1. Which journals cover the world's literature about _____ (art, physics, or whatever)?
2. In which countries are these journals published?
3. In what languages is the literature about _____ written?
4. How much is written in each subject field of _____, where is it published, and in what language?

In the above series of questions, 1, 2, and 3 are equally important for

information scholars and specialists, while question 4 has value primarily for the specialist and secondarily for the information scholar. Investigation of all the questions is possible in numerous narrowly or broadly specialized areas such as physics,[6] biology of mammals,[7] or cardiovascular medicine.[8] It would be a meritorious task to undertake as complete as possible a bibliographical listing of these studies. Arranged according to specialized areas, an overview would be obtained which would benefit many budding information projects, e.g., those ranging from one involving equipping libraries to one designed to find the best way to build up an information service. In both of the cases mentioned above, calculation of the active life-span of scholarly literature should also be mentioned as an important consideration.[9] This factor often causes a limitation of the literature collection in terms of time-span, while the application of Bradford's law permits a limitation of the range by supplying the essential core journals.[10] These two factors may also be deduced by analyses of bibliographies, as discussed in Brookes's paper in this issue.

If question 4 above is extended somewhat further, one can also convincingly distinguish trends in specialized areas. Methodical analysis of the specialized bibliographies may also be applied here.[11] An unpub-

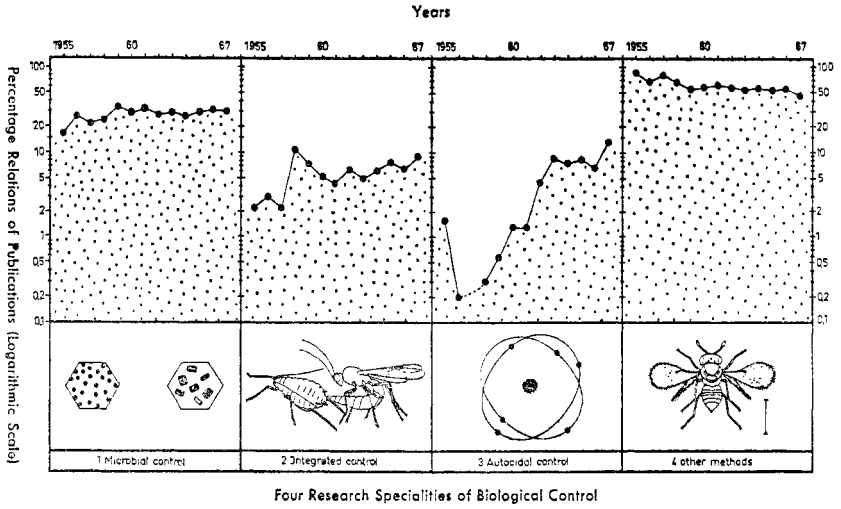

FIG. 1. Percentage Relations of Publications in Four Research Specialties of Biological Control. Data taken from "Bibliography on Biological Control," which appeared in the journal *Entomophaga* (Paris).

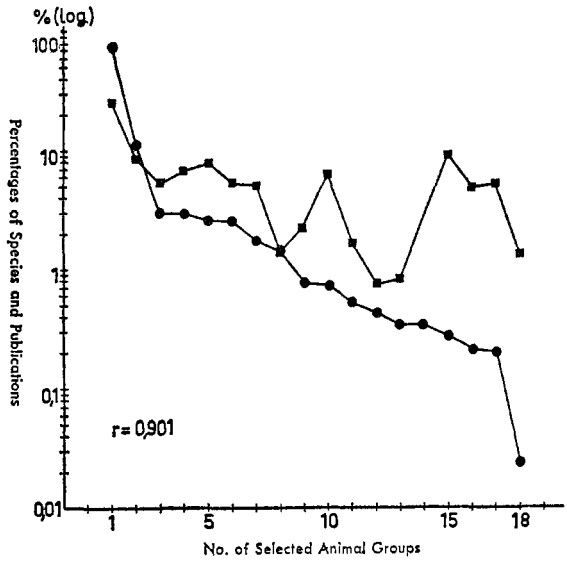

FIG. 2. Relationship between the Number of Species to Publications in Eighteen Animal Groups. Species are signified as circles, publications as squares. Data are taken from *Zoological Record* (London) issues from 1955-1967.

lished example from my own current studies can demonstrate this in the field of biological control of pests (see figure 1). First of all the information flow is represented purely descriptively in one area of research.

In addition to determining how much, where, and what, it is also important to determine the factors which are responsible for course and strength of the information flow. Some very promising formulations in this area have been furnished by de Solla Price.[12] A comparison to this kind of investigation at the international economic level would be the gross national product. Correlations with journals in the "environmental protection" sector are thereby clearly revealed.[13]

In the field of biology the number of species within systematic categories are good comparative parameters for the corresponding share of literature.[14] An example of this is represented in figure 2, which shows the highly significant positive correlation of 0.901 for the numbers of species and the corresponding share of publication for 18 selected animal groups (for details see reference 14).

Introduction

Citation indexes have been a major aid in tracing tendencies and sociological and historical interdependencies in various areas of research.[15] Although it has been found preferable first to investigate the sciences, the *Social Sciences Citation Index,* scheduled to be introduced in 1973, should also provide an improved basis for corresponding studies in the arts and the humanities.

Even the few examples given in this short summary should show that the analysis of bibliographies is a method which, from the viewpoint of information studies, deserves urgent priority in new projects. Analyses of bibliographies can give valuable insights and overviews not only to the information specialist himself, but also to the specialized researcher. This issue of *Library Trends* is intended to be a preliminary, useful overview of the subject and will possibly give an incentive to further studies in the field of "analyses of bibliographies."

References

1. Malin, Morton V. "The *Science Citation Index*®: A New Concept in Indexing," *Library Trends,* 16:374-87, Jan. 1968.
2. Lancaster, Frederick W. "The Evaluation of Published Indexes and Abstract Journals: Criteria and Possible Procedures," *Medical Library Association Bulletin,* 59:479-94, July 1971.
3. *Ibid.,* p. 484.
4. Simon, H. R. "Zur Analyse Biologischer Bibliographien," *Heidelberger Jahrbücher,* 15:111-33, 1971.
5. Wyatt, H. V. "When Does Information Become Knowledge?" *Nature,* 235: 86-89, Jan. 14, 1972.
6. Kennan, Stella, and Atherton, Pauline. "The Journal Literature of Physics: A Comprehensive Study Based on Physics Abstracts." New York, American Institute of Physics, 1964.
7. Anderson, Sydney, and Van Gelder, Richard G. "The History and Status of the Literature of Mammalogy," *BioScience,* 20:949-57, Sept. 1, 1970.
8. Frick, Barbara F., and Ginski, J. M. "Cardiovascular Serial Literature: Characteristics, Productive Journals, and Abstracting/Indexing Coverage," *Journal of the American Society for Information Science,* 21:338-44, Sept. 1970.
9. Line, Maurice B. "The 'Half-Life' of Periodical Literature: Apparent and Real Obsolescence," *Journal of Documentation,* 26:46-54, March 1970.
10. Goffman, William, and Morris, Thomas G. "Bradford's Law and Library Acquisitions," *Nature,* 226:922-23, June 6, 1970.
11. Rothman, Harry, and Woodhead, Michael. "The Use of Citation Counting to Identify Research Trends," *Journal of Documentation,* 27:287-94, Dec. 1971.
12. de Solla Price, Derek J. "Measuring the Size of Science," *Proceedings of the Israel Academy of Sciences and Humanities,* 4:98-111, 1971.
13. Simon, H. R. "Informationsmöglichkeiten zum Umweltschutz," *Natur und Landschaft,* 47:252-55, Sept. 1972.

14. ———. "Animal Species and Publications," *Journal of the Washington Academy of Sciences.* (In press)

15. Weinstock, Melvin. "Citation Indexes." *In* Allen Kent and Harold Lancour, eds. *Encyclopedia of Library and Information Science.* Vol. 5. New York, Marcel Dekker, 1971, pp. 16-40; and Garfield, E. "Citation Indexing for Studying Science," *Nature,* 227:669-71, Aug. 15, 1970.

Secondary Services and the Rising Tide of Paper

J. MARTYN

SOME YEARS before the phrase "information explosion" became current, the author of the Book of Ecclesiastes observed that "Of making many books there is no end" and, voicing a view that one suspects is shared by many of those for whom information services are hopefully designed, added that "Much learning is a weariness of the flesh." Since that time, the journal literature has appeared, and the potential weariness of the flesh has been correspondingly increased. Strictly speaking, Ecclesiastes has fallen into the common trap of equating books with knowledge because today one still speaks of an information explosion when what is properly meant is a paper explosion; that is to say, there is a tendency to assume that because the number of published items is continually increasing, the fund of human knowledge is being added to by a similar amount. The distinction between information and documents is an important one, because what abstracting and indexing services are basically concerned with are pieces of paper, and the fact that there are a great many of them causes most of the problems of secondary services.

At various times the number of papers published annually has been estimated at anything up to a million and a half. Variations in estimates arise because of differences in the bases examined, in the definition of what constitutes a published paper, and in the selection of the area of knowledge examined. Most estimates concern themselves with science and technology, and in this field the estimate published by Vickery in 1968 is probably the most firmly based.[1] He defined a paper as an authored contribution to a scientific or technical journal, and, by careful sampling of current issues of journals held by the National Lending Library, arrived at an estimate of approximately 850,000 papers pub-

J. Martyn is a Research Scientist at the Aslib Department of Research and Development, London.

lished annually. Comparing this figure with other estimates, it seems reasonable to suppose that by the end of the 1960s the annual output of papers in science and technology was of the order of 1 million.

A more recent study by Vickery[2] of the growth of journal literature, from its first appearance to the present, suggests an overall approximate doubling period of twenty years, and a cumulative total of 30 million papers. On these estimates as foundation one can rear an imposing edifice of entertaining speculation. For example, the British Museum having between seventy and eighty miles of shelving, it would be possible to house one copy of every scientific and technical paper ever published on the existing shelves, if they were empty. As there would be some room to spare, it follows that the volume of books at present housed there is greater than the volume of papers that currently exists. Or, to put it another way, there are more words in books than there are in journals. Considering the rate at which the British Museum adds new shelving, it is likely that the doubling period of the book literature is about twenty-three years, not grossly dissimilar to that of the journal literature. It follows that not only are there more words in books than in journals, but there always will be, and the gap between the two is widening.

As it is clearly harder to read all the books than to read all the journal papers, why not abstract books, rather than papers? Of course, the foregoing flight of fancy has neglected the journal literature of social science and the humanities, as one often does, and has glossed over the proportion of the British Museum library that is fiction, social science or humanities, or is simple translation of other volumes, and tends also to ignore the higher repetition rate of information in books as compared with journal papers. Nevertheless, it was an example of the sort of extrapolation one can (and frequently does) engage in in the information field, when presented with a little nugget of near-fact.

Vickery's estimate of 30 million papers referred to science, technology and medicine only. It is difficult to extend the estimate to the social sciences and the humanities, partly because of the lack of comprehensive listing, partly because of difficulties in defining social science (or humanity), and partly because the more literary a document becomes, the less likely is it to be classed by an information scientist as literature; that is to say, the scientific paper is an inherently more definable unit act of communication in science than the vehicles of communication appropriate to other scholarly activities.

There are a great many secondary services whose purpose it is to

provide references to, and frequently abstracts of, portions of the total mass of literature, often in an indexed or classified form. The portions chosen can be either topic or discipline oriented like *Chemical Abstracts,* mission oriented like *Food Science and Technology Abstracts,* quality oriented like the *Science Citation Index,* or with a national orientation like *British Technology Index.* The coverage policy can be selective, usually on a quality basis, or comprehensive. The total number of secondary services is uncertain. In 1963 the National Federation of Science Abstracting and Indexing Services published a listing of all the science-technology services publishing more than 100 abstracts a year that they had been able to identify, and this listing contained 1,855 entries;[3] NFSAIS is currently working on a revision of this list. Separate sections of large services, such as the parts of *Referativny Zhurnal,* were shown as individual entries, and a number of the services listed were abstracts sections of otherwise normal journals. Some services were available in several formats, e.g., printed codex or cards. There is no doubt that the forthcoming revised list will show considerable variation from the 1963 list, particularly in terms of the variety of formats available, since there are now in the region of 100 services available on magnetic tape.

In the intervening period there has also been an accelerated tendency for large services to subdivide or to provide varieties of bibliographic output, as for example *Chemical Abstracts Condensates, Chemical Titles, Chemical-Biological Activities,* and so on. And undoubtedly the numbers of available services have increased. Each individual service will also certainly show an increase in the numbers of items published annually. *Chemical Abstracts* has jumped from 269,293 in 1967 to 350,105 in 1971, and in the same period INSPEC has jumped from 71,032 to 148,944 and *Engineering Index* from 51,670 to 85,243. Other services will have increased similarly, partly as a response to increases in the primary literature, but also because of improved performance in acquiring the relevant literature, and especially because of policy decisions whose effect is to widen the scope of the secondary services by including much material which might previously have been regarded as peripheral.

Vickery, in the growth study referred to above,[2] also points out that an examination of the growth of six major abstracting services over the past decade shows a doubling period of six and a half years, which is a faster rate of increase than that of the primary literature itself. A recent study of information on tape,[4] carried out by the Aslib Research and

Development Department, identified about seventy services for which it was possible to estimate the total number of entries made up to December 1970, and the annual number of entries. This study indicated that at least 3.5 million machine-readable entries are now recorded annually, and the total number of entries on tape was, at the end of 1972, about 24 million.

It is already obvious that there are many more entries made annually in secondary services than there are primary papers published. Since no service is notifying imaginary references, it follows that there is very considerable overlap among the secondary services, with multiple notification of many items. It is worth noting that the great majority of services examined in the Aslib tape study are in the English language, so there is no evidence to support the hypothesis that the overlap is caused by abstracting or notification in different languages: it would not be unreasonable to assume that, had it been possible to include non-English language services in the study, a higher rate of duplicate notification would have been suggested.

Martyn[5] noted that the study of a number of reasonably comprehensive specialized bibliographies showed that about one-fifth of the references contained in the bibliographies had not been picked up by any major secondary service, and concluded, with the specialized bibliographies used being drawn from a number of areas of science and technology, that it was not unfair to assume that, over the whole field, a similar situation prevailed. A more recent unpublished work by Martyn, in which items notified by the Biodeterioration Information Centre have been sought in Institute for Scientific Information coverage, *Chemical Abstracts* services, *Biological Abstracts, Index Medicus* and *Food Science and Technology Abstracts* (all services which may be searched by computer), has confirmed the original view because, of the 1,874 references to 1969 journal literature sought, more than 23 percent proved not to have been covered by the services examined.

We now have the picture that, so far as science and technology are concerned, roughly 1 million items of primary journal literature appear each year. Probably (augmenting the Aslib tape coverage figure to allow for services not yet on tape) more than 4 million entries appear annually in secondary services. As approximately one-fifth of the items published are not notified, it is reasonable to conclude that, of those items which *are* notified, each is notified, on the average, five times. It may further be assumed that many of the notified items are notified many more than five times. This is a result of the natural evolution of

the system of coverage of the literature of science and technology. The major services, which are mainly the longest established, cover the literatures of specific disciplines; coverage of their core literatures is generally complete, or nearly so, but coverage tends to fall away from completeness in the fringe areas, as is natural. The core of a discipline is clearly defined, and the relevance of a document to that core can be easily assessed, but a fringe area is by its nature less clearly defined, and decisions of relevance become more subjective. Usually a major service is operated by full-time professionals, with paid abstractors and indexers, and a marketing system organized along commercial lines. Exchange of appropriate material among services has been common for a long time, and is gradually becoming more systematic.

Together with the major services there is a population of smaller secondary services which cover subdisciplines, whose coverage is often wholly contained within the major services, but which continue to exist either because they provide their customers with more specific information (e.g., the user has to buy less of what he does not need in order to get what he does need), or because the quality of their product is superior by virtue of being better indexed or having more information-content per item notified. These services also tend to be reasonably long-standing and professionally organized, but not infrequently rely on voluntary or semivoluntary (i.e., underpaid) assistance with their input.

There are also a large number of small services which have come into being in order to cover interdisciplinary areas, the areas which lie between the defined coverage areas of the bigger services or in some cases outside them. Being interdisciplinary, they tend to serve developing areas of science which are relatively novel, and, therefore, often of no great age. Because the areas of interest are novel, there is often not a very large population of users to support them, so their potential income is small and they rely on volunteer labor, and the absorption of their overheads by large organizations such as, for example, universities which provide accommodation and facilities, or major libraries which allow free access to their journal stocks, to keep their apparent costs down. Some are run by professional information workers and some are not. In many cases the marketing of their products is far from perfect. Where diversity can exist, it does.

Some services attempt to cover their chosen area of interest comprehensively and include everything relevant which can be discovered by scanning a wide range of journals, like *Chemical Abstracts* which "in

1971 . . . combed the contents of some 12,000 scientific and technical journals published in 56 different languages and chemical patents issued by 26 nations." Some restrict their coverage to journals alone, or patents or the report literature. Some cover selectively, imposing a degree of quality control on the material accepted for inclusion. Some simply claim to cover a field as well as they are able by drawing from a defined set of sources, like the *Bibliography of Reproduction* which claimed an approximate 50 percent coverage of the field of mammalian reproduction. ISI, unlike the majority, covers a stated list of journals comprehensively, including everything published in the covered journals. The quantity of information presented varies from a bare reference to a detailed informative abstract, often acceptable as a substitute for the original document. The style and quality of indexing varies similarly, and the time elapsed between publication of a document and its notification can vary from two or three weeks to, in extreme cases, several years.

Just as the primary literature grows, develops, overlaps, multiplies and occasionally dies with little to check it but the pressures of the market (moderated sometimes by direct subsidies, or the unrealized subsidies that arise from defective accounting), so the secondary literature exhibits the richness, the vigor of growth, the variety, the extravagance, the untidiness and the illogicality characteristic of the majority of human activities in a free society. Irrationality is the price of freedom in more than a purely philosophical sense, and even in those systems where sanctions can be imposed to enforce order, humanity, like cheerfulness, keeps breaking through, and the systems therefore fall away from perfection.

What the effects of this lack of systems are on the user depends very much on the nature of the user. If his interests lie within the boundaries of a long-established scientific discipline, he is liable to find the present system adequate to his needs, particularly if his requirement is for less than total coverage of his topic of interest. As total coverage is liable to produce references to many more documents than he is willing or able to read, he is predisposed to accept partial coverage, although welcoming some assurance that what he misses is of less importance than what he finds. If, however, as is perhaps more commonly the case, his immediate requirement is for a piece of information whose normal habitat is in a discipline outside his own, or he requires a body of references which are scattered among the literatures of several disciplines, or which are even difficult to locate within the normal framework of

the traditional disciplines—in other words, if his temporary or enduring need is for interdisciplinary information—he faces serious problems. The first problem is the choice of service to search, which is largely a matter of trial and error. Having selected a service, if he is seeking a reasonably comprehensive or, better, representative, sample of literature on his interdisciplinary topic, he can be certain that all his needs will not be met from within that one service alone. He must therefore turn to other services, but if he does this, he can be certain that not only will he be unlikely ever to achieve anything approaching comprehensive coverage, but also he will inevitably find the same items over and over again.

The first problem may or may not be serious. There are some grounds for supposing that the information that is not picked up by the secondary services tends to be of a lower level than that which is picked up; lower in the sense of generally not being the first appearance of the contained information, but rather rewritten versions of it produced for scientists in different fields or technologists in the same field as the field of origin of the original information to which it refers.

The second problem is more important, particularly in the case of an individual searching services based on machine-readable records, because in the majority of cases the payment for records retrieved from these services is directly related to the number of items retrieved or expected to be retrieved, and while the user population may become resigned to the prospect of paying for a number of unwanted references inevitably retrieved with those which are wanted, it is less likely to accept the necessity of paying afresh for items which it has already bought, however relevant they may be.

What is the purpose of the system of secondary services? It is seldom stated in a global context, being usually related to a particular population. Any specific secondary service is generally aimed at covering the literature which is likely to be of use or interest to scientists or others in a stated discipline or subdiscipline, or with a stated mission; it tends, that is to say, to be user-oriented. This unfortunately means that while there are very many users whose average needs are satisfactorily met by existing single services, there are also, inevitably, many other users, either in interdisciplinary areas or in some areas of applied science and technology, whose needs are not met, just as a normally well-served user occasionally has an unusual information need which is difficult to meet from his habitual sources. In an ideal world, access to all parts of the primary literature should be equally possible for all potential users,

and to afford such access should be the object of the system of secondary services as a whole.

To insure equal access to all documents, it is necessary to insure that references to all documents exist somewhere in the overall system. So far as the book literature is concerned, this is gradually being achieved through the agency of the various national bibliographies, and when all countries are eventually producing MARC-like services for their own book output, then, errors and omissions excepted, world coverage of monograph material will be complete. It should be noted that the existence of national bibliographies does not preclude the existence of specialized bibliographies, but rather facilitates their production.

National bibliographies work because they are organized on the basis of the physical properties of the items they cover; they are identifiable as books, according to an agreed definition, and they are published in identified locations. Discipline- or mission-oriented secondary services are less successful in terms of coverage because they are based on the intellectual properties of the documents they cover—what they are about, rather than where they come from. A "journal-paper MARC" operated on a national basis is as feasible as existing MARC systems, and would at least provide comprehensive coverage, searchable by journal title, article title and author, of the journal literature of the individual countries operating such a system.

Such a system would take some time to implement, and might well require legislation, if only to provide enabling funds. It would inevitably be a considerable time before the world's literature could be covered in this way. Attempts to approximate a consolidated coverage of the journal literature could be made if it were practicable to merge the existing major tape-based secondary services in such a way as to create a unified file containing the author, title and reference of each item contained in the major services, and perhaps augment it with additional material known not to be covered. One problem which would have to be faced would be identification of noncovered material, and this would be greatly facilitated if the major services could state categorically which journals they cover completely. (This would be of value immediately, in that searchers of more than one service could use a logical NOT to prevent re-retrieval of at least some of the items found in the first search.)

Another problem would be the elimination of duplicate references, a problem which is currently intensified by the great variation in timelag of available systems. A reference appearing in service A perhaps a

month after first publication may appear in service B after a further three months, and again in service C six months later still; this problem can be met by using a journal-item-received register file and automatically deleting from each fresh input tape every item already shown to be held, but it is a cumbrous and unwelcome necessary addition to a notional system. The result, however, would be a single source which could be searched by a user without incurring cash penalities for duplicate retrieval.

There are basically three routes towards a unified, controlled system of secondary services. They are to operate the system yourself, with or without using the elements already existing, to coerce others into operating the system, or to persuade them to do so. The first route is essentially what has been suggested here. Coercion is not only morally undesirable, but also not practicable, so may be dismissed. Persuading others, usually known as "international cooperation," is the route which has been and is most often followed, not without a certain amount of success, but anyone with any experience of the international information field knows this is an arduous route, bedeviled with questions of national, organizational and individual self-interest. One recognizes that in the information field, as in other perhaps more important or more urgent problem areas, international cooperation is the only feasible long-term solution, but it is suggested that until this happy state be reached, some stop-gap operations would not be unuseful to the actual users of information, for whom, after all, the secondary services are notionally intended, and who, ultimately, pay the bill.

References

1. Vickery, B. C. "Statistics of Scientific and Technical Articles," *Journal of Documentation*, 24:192-95, Sept. 1968.

2. ———. *The Growth of Scientific Literature.* (To be published)

3. National Federation of Science Abstracting and Indexing Services. *A Guide to the World's Abstracting and Indexing Services in Science and Technology* (Report no. 102). Washington, D.C., NFSAIS, 1963.

4. "Information on Tape." Aslib Research & Development Department, Aslib, 1971. (Internal report)

5. Martyn, John. "Tests on Abstracts Journals: Coverage, Overlap and Indexing," *Journal of Documentation*, 23:45-70, March 1967.

Numerical Methods of Bibliographic Analysis

B. C. BROOKES

IT IS ONLY in the last eight or ten years that the numerical aspects of bibliography have attracted attention, although some of the numerical regularities that occur in bibliography have been known for thirty or forty years. Results are, therefore, still meager and applications are still few. Moreover, most of the work so far reported has been limited to numerical analysis of the literature of the natural sciences. This is in part because the secondary sources in the natural sciences are the best organized and so provide the most accessible data; in part because the literature of the natural sciences are the least restricted by linguistic barriers; and in part because the proposed world-wide systems, such as those advocated in the UNISIST report, offer an immediate field of application in the design of economic and efficient systems based on the results of numerical bibliographic analysis. However, the field of possible application is gradually widening: serious efforts are now being made to organize the more diffused literatures of the social sciences, for example.

The main practical purposes that numerical analysis can serve are based on the belief that quantification is a necessary component of the design of economic information systems and that measurement of the key processes of an information system is a necessary component of management control. The present main objectives of numerical analysis are:

1. the design of more economic information systems and networks;
2. the improvement of the efficiencies of information-handling processes;
3. the identification and measurement of deficiencies in present bibliographical services;
4. the prediction of publishing trends; and
5. the discovery and elucidation of empirical laws which could form the basis for developing a theory of information science.

B. C. Brookes is Reader in Information Science in the School of Library Archives and Information Studies, University College of London, England.

Estimating the Completeness of a Bibliography

Strictly speaking, no bibliography can ever be *proved* to be complete. The bibliographer can only publish his final list and so, in effect, challenge all who may be interested to point to any omission they may note. The bibliographer who strives to prepare a comprehensive bibliography will usually turn first to the most accessible and productive sources contributing to his subject. But as his work progresses he has to seek and identify other less productive sources. The search, ever more penetrating and wide-ranging, is continued while relevant items are found. But new finds occur with steadily decreasing frequency in spite of continued effort. There is no positive signal which indicates to the bibliographer when his search has been completed; the only sign is the absence of further relevant items.

However, the law first formulated by S. C. Bradford,[1] and known by his name, offers the possibility of estimating the number of sources and the number of items that one can expect to find. This estimate is based only on knowledge of a small but sufficient number of the most productive sources. Unfortunately, Bradford formulated his "law of scatter" in two versions which Wilkinson[2] has recently shown to be formally different although closely similar. But both formulations lend themselves to methods of estimating the size of a comprehensive bibliography if the subject and the range in time are first well-defined.

In its original form, Bradford's law said nothing about comprehensiveness. But obviously a bibliography must be finite and the number of items produced by the least productive sources cannot be less than one. When this consideration is expressed in one of the two formulations of Bradford's law, a simple graphical technique for estimating the size of the complete bibliography can be devised. It requires the drawing of the "bibliograph."[3]

The most productive sources are first ranked in decreasing order of productivity. The cumulative sums of items found from these sources are then plotted on a graph as shown in figure 1. The most convenient graph paper for this purpose is semilogarithmic; the linear scale is applied to the cumulative sums and the logarithmic scale (which, although marked 1, 2, 3 . . . 10, 11, 12 . . . in successive digits, actually spaces them according to the logarithms of the digits) is used to indicate the ranks of the sources. So, on the graph, point A indicates the number of items yielded by the most productive source, point B indicates the number of items yielded by the two most productive sources together, and so on. The first few points will be found to lie on a rising curve which, sooner

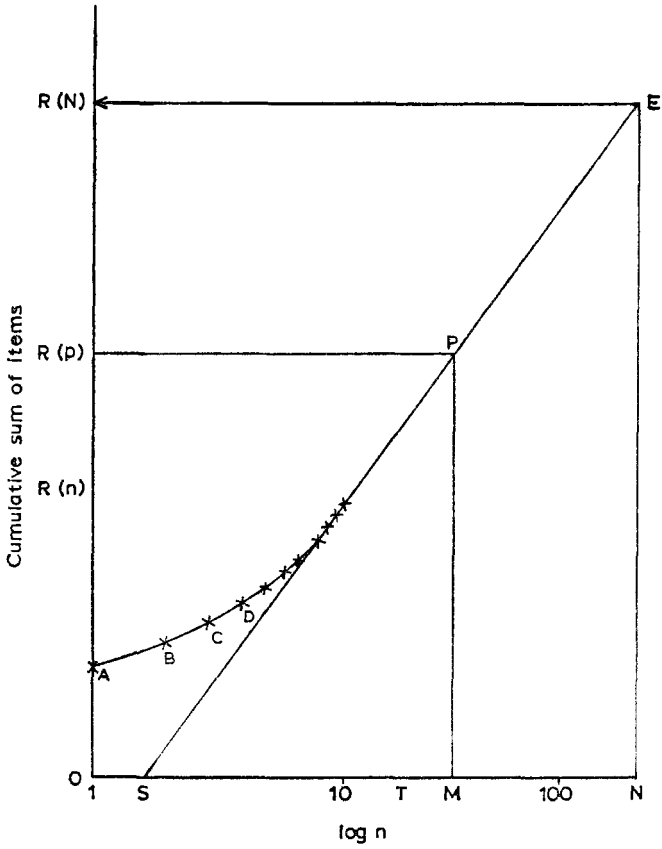

FIG. 1. Plotting the Bibliograph

or later, will run into a straight line. As soon as the straight line is definite enough, an estimate of the end-point of the line can be made.

The straight line is extended to meet the axis of log n (at S in figure 1) and some convenient point P is marked on the straight line. It can then be shown that the total number of sources, N, expected to contribute to the bibliography is given by

$$N = \frac{R(p) \cdot OT}{3\,SM} \tag{1}$$

where $R(p)$ is the number of items corresponding to the point P and where the lengths OT and SN are measured as accurately as possible in millimeters.

When N is known, it is possible to estimate the total number of items to be expected. It may be possible to mark the point N on the graph and so to mark the corresponding point E on the continuation of straight line SP. The required value of $R(N)$ can then be read from the scale of $R(n)$ on the left hand side. Alternatively, $R(N)$ can be calculated from the formula

$$R(N) = N \log_e N/S \qquad (2)$$

where it is the number corresponding to S on the log n scale. (A table of "natural" logarithms is needed.)

How realistic is the estimate? It is not possible to *prove* that E *must* be the end-point of the line because there is no logical reason why a comprehensive bibliography should conform so precisely to a mathematical law. Yet the technique has now been tested many times, especially against computer-produced bibliographies derived from retrospective searches of large data bases such as MEDLARS, and seems to be realistic. The advantage of using computer-produced bibliographies of this kind is that the items found must all conform to the search question as formulated for the computer search program. The relevance to the subject specified by the question is therefore uniformly controlled and it does not matter (for the purpose of testing the technique) whether the search question is "correctly" formulated or not. In every such case real and predicted end-points are very close indeed.

When the technique is applied to manually produced bibliographies the real and the estimated end-points can differ appreciably. A common occurrence is illustrated in figure 2. Here, if the plotting of the graph is continued beyond the points required to determine the straight line, the plotted points may fail to maintain the linear climb and fall away in a droop to end at some point such as G. In such cases it is plausible to argue that the bibliography is not complete in the sense in which the technique requires. For example, the bibliographer may quite reasonably state that he has been selective. He may say that he has noted only those items which are "of professional interest" or that he has omitted exact translations of some items into other languages. But, in our own experience, wherever a droop has been observed, it has always been possible to indicate either some selectivity or some omissions.

It has been noted, however, that the point E as calculated slightly overestimates the total number of items, though not the total number of sources. This fault arises because, as the graph nears the end-point, the sources end with a number which provides three items, a larger number

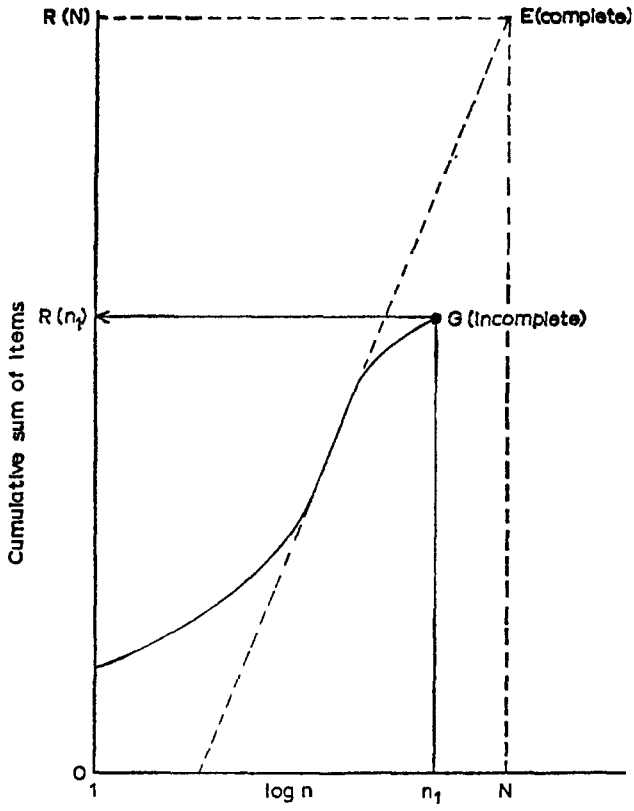

FIG. 2. The 'Droop': An Indication of Incompleteness

which provides two items and a still larger number of sources which provide only one item each. When the corresponding cumulative sums are plotted on a logarithmic scale they do not lie exactly on the straight line but form a series of lengthening arcs which intersect on the straight line (figure 3). The last one, however, is open at the end-point. When the complete data are carefully plotted, the graph ends in an open hook which ends slightly below the estimated end-point E. The "hook" however is clearly distinguishable from the "droop" mentioned earlier: the hook is concave upwards whereas the droop is concave downwards.

With these reservations Bradford's law can be used to provide reliable estimates—but of what precisely? The user of this technique has to appreciate that the end-point is determined by the items he includes in the data which initiate the curve and straight line and that these data

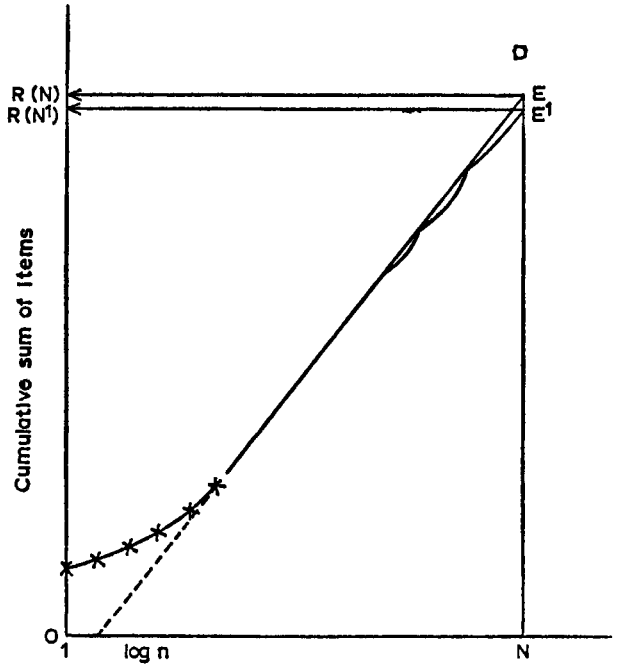

FIG. 3. The Arcs Near the End of the Bradford Linearity

define the totality he is estimating. For example, if it happens that these initial data are all derived from the most productive sources of, e.g., items published only in the *English* language, then the totality will be of items published in the English language. If the initiating data all refer to items published in, e.g., 1972 in any language, then the totality will be of items published in 1972 in any language. Most of the faults about the application of the technique that have been reported arise from lack of appreciation of the fact that the initiating data must be bibliographically well-defined in subject relevance and in period of publication and also be complete and exact as far as they need to go. And the reliability of the estimate is naturally critically dependent on the precision of the initiating data.

OTHER OCCURRENCES OF BRADFORD'S LAW OF SCATTER

The mechanism underlying the generation of a Bradford distribution seems to consist of two competing processes. The primary process is a

bandwagon effect: in bibliographical terms, authors would in general prefer to publish their papers in the core journals which correspond to the sources on the curved part of the bibliograph. But these journals receive more good papers on the particular subject than they can or wish to publish. So their standards of acceptance are high and this secondary restrictive process then pushes some papers out into the peripheral journals. The Bradford distribution follows: it represents an overall bandwagon effect which is modified by a restrictive effect over the sources at the core.

If this explanation were valid, the Bradford distribution should arise in other situations in which similar processes occur. It does. The items borrowed from a library follow the same law: the restriction on borrowing occurs because there are always a few items in such demand that some keen users have to wait.[4] The users too, ranked in order of the

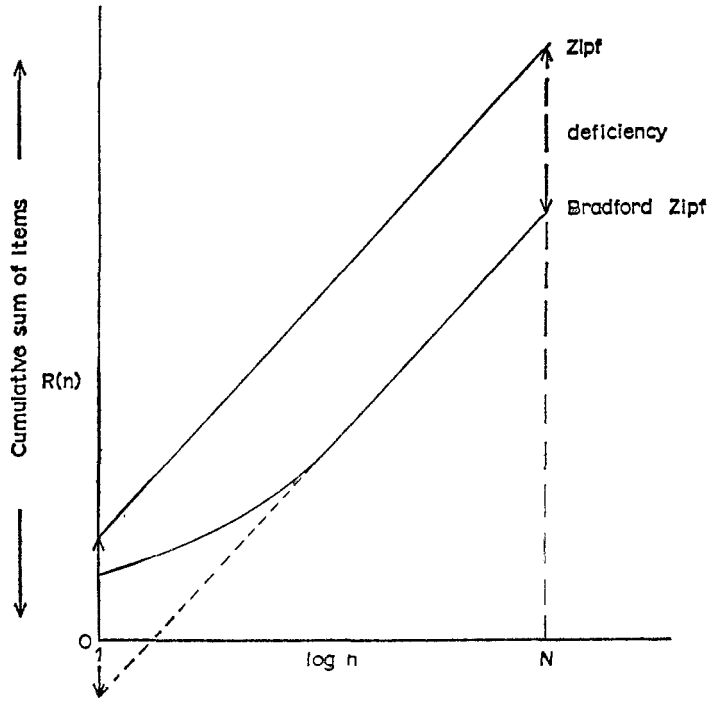

FIG. 4. Cumulative Zipf (Unrestricted) and Bradford-Zipf (Partially Restricted) Distribution

number of items they borrow, form another Bradford set: keen users compete with each other for the core items of the library and so are sometimes disappointed not to find the specific items they seek. It has been suggested (not too seriously) that a library could most economically be designed to provide the best possible service for all its users simply by striving to meet, as far as possible, the demands of its identified core users, because all other users will be content with what is so provided.

It has also been noticed that at a conference the contributors to the discussions, ranked according to the frequency of their questions or contributions, yield a Bradford set of data. Here, competition arises among the would-be contributors to "catch the chairman's eye," and the good chairman usually prefers to call a new contributor to the discussion when the chance arises. The index terms assigned to documents also follow a Bradford distribution because those terms most frequently assigned become less and less specific and therefore increasingly ineffective in retrieval. The publication of books by publishers follows the same law.

On the other hand, when there is no competition or other form of restriction, there is no core. At the present time, authors of scientific papers are generally allowed to cite as many references as they wish. The data which relate to items cited yield a bibliograph which consists of a straight line with no initial rising curve (figure 4). This is a cumulative Zipf distribution corresponding to the linear nonrestrictive part of the Bradford-Zipf distribution.

Applications of Bradford's Law

Comprehensiveness of abstracts and other bibliographic services

Every specialized interest in the natural sciences is now provided with abstracts services and some with retrospective search of cumulated data bases. These services are usually presented implicitly as though they covered their particular fields comprehensively, but it is well to check the coverage offered. The bibliograph can be revealing. If the graph ends in a droop it suggests that the service provided is not comprehensive, but before reaching such a decision the analyst must insure that his data are well-defined and are exact. For example, if it were decided to check the completeness of the list of items published in 1972, then adequate time must be allowed to permit the inclusion of all such items in the service being checked.

SELECTION OF SOURCES

The complete coverage of any well-defined subject is a costly objective, so most librarians are faced with the problem of deciding how to make optimum use of limited resources. If, for example, the attempt is made to identify the sources that contributed to the Bradford-estimated totality of some well-defined subject during the year 1970, it will be found that there may be 100 sources which provided only one item. If a similar analysis is made a year later, again it may be found that about 100 sources have yielded only one item each. But the two lists of one-item sources will not be identical: in each list there will be found sources which do not occur in the other. In short, whatever the subject may be, there is a very wide fringe of sources which contribute less frequently than one item per year, and continuous search is needed in this peripheral area to discover these occasional items. There is no discernible limit to comprehensiveness in this wide sense.

The librarian, therefore, has to be selective. The analysis of sources required to produce the data for the bibliograph requires the sources to be ranked in order of decreasing productivity. For a given annual expenditure on the acquisition of sources, it therefore provides the data needed to make selection a rational process, and the bibliograph enables the librarian to estimate the percentage coverage he can afford.

The core sources together always constitute a large fraction of the Bradford totality at relatively low cost and, because they also represent the most active sources, they can be regarded as providing the essential working minimum for any special library. Beyond this nuclear zone of the most productive sources there are sources which are gradually less and less productive. Moreover, there is some evidence (not yet fully established) that the nuclear zone produces the most frequently cited items and therefore, presumably, the most significant also. The general Bradford formula makes it possible to estimate how many of the most productive sources would yield any specified fraction p of the total number of items. The formula is

$$R(n) = N \log n/s \ (1 \leq n \leq N) \tag{3}$$

where

$R(n)$ = cumulative total of items contributed by the sources of rank 1 to n,
N = total number of contributing sources,
s = a constant characteristic of the literature.

Then

$$R(N) = N \log N/s \tag{4}$$

is the total number of items contributed by the N sources.

As a simple numerical example, consider a scientific literature which produces a total of 2000 items from 400 sources. We then have, substituting in (4),

$$2000 = 400 \log 400/s$$

so that

$$\begin{aligned} \log s &= \log 400 - 5.0 \\ &= 6.0 - 5.0 \quad \text{(natural logarithms)} \\ &= 1.0. \end{aligned}$$

It is now possible to complete the calculation, since

$$p = \frac{R(n)}{R(N)} = \frac{\log n/s}{\log N/s}$$

which on solving for $\log n$, yields

$$\log n = p \log N + (1 - p) \log s. \tag{5}$$

For the particular values of the numerical example, we have

$$\begin{aligned} \log n &= p \log 400 + (1 - p)\, 1.0 \\ &= 6.0\, p + (1 - p) \\ &= 5.0\, p + 1. \end{aligned}$$

If we now put $p = 0.8$, 0.6, 0.4 and 0.2 successively, the corresponding values of $\log n$ are 5.0, 4.0, 3.0 and 2.0 successively. And referring to tables of natural logarithms or exponentials, we find values of n as shown in the following table. Note in table 1 that the 20 most productive of the 400 sources contribute 40 percent of the total number of items.

TABLE 1

SOURCE COVERAGE BY PERCENTAGE

Percentage Coverage of Items	No. of Sources (n)	Percentage of Sources
100	400	100
80	148	37
60	55	14
40	20	5
20	8	2

The same estimates can also be found graphically. On semilogarithmic paper, with the rank n marked along the log scale and $R(n)$ along the linear scale, identify the two points: $R(N) = 2000$ with $N = 400$ and $R(n) = 0$ when $n = s = 2.7$. The line joining these points enables any corresponding values of n and p to be read directly from the graph.

These measures are, of course, purely quantitative; so it would be reasonable to consider sources on both sides of the proposed cut-off to insure, for example, that a source in a language unfamiliar to the users, or relatively costly, difficult to acquire or subject to delay is not retained in favor of another source, of slightly less productivity, in a familiar language, less costly or easier to acquire. The optimum cut-off has to be applied with judgment based on bibliographic knowledge of the subject area.

Any such cut-off leaves the librarian with the problem of searching for significant items provided by sources which he has decided not to acquire. Notice of such items should be given by the abstracts services or other secondary sources and their mode of acquisition will depend on the local circumstances. In Britain, the National Lending Library strives to acquire copies of all scientific sources and offers a photocopy by return mail of any item that is bibliographically well described. In such a case, the optimum cut-off point can be determined by estimating the cost, in terms of the delay to the users as well as of the actual monetary cost, and comparing this figure with the cost of acquiring items from the ranked sources as the rank (and therefore the cost per item acquired) runs from 1, 2, 3 . . . onwards until the two costs are equalized.[5]

This technique of determining the economic holdings of scientific libraries can be applied to any hierarchical system of libraries which consists of a number of small local libraries with special interests, a smaller number of regional libraries which service groups of special libraries, and a national library which offers the most comprehensive coverage. The conditions under which the regional libraries become redundant can be explored.[6]

The Measurement of Obsolescence

The world's knowledge of pure science is embodied in its published literature which is in part a description of theories and data that are wholly new, in part an updated version of earlier theories and a reinterpretation of older data in terms of the newer theories. So, at any given time, the current scientific literature is in part new and in part retrospective although the range of the retrospective view varies from one

science to another and even within a science, as in astronomy for example, there may be wide differences as there are between descriptive astronomy and astrophysics.

If a library's resources are limited and its users are primarily concerned with current research, there is little point in retaining in perpetuity all the back numbers of all the journals it takes. The usage of back numbers declines with age but the little-used volumes demand constant shelf space and maintenance services which add to the overall library costs. It is possible to estimate the optimum cut-off graphically when the aging rate is known, and economical to apply it when copies of any papers called for from the discarded volumes are readily available elsewhere.[6]

However, in measuring the rate of obsolescence of a literature we have to distinguish between the usage of the literature by those who contribute to it (who are its primary users) and the usage of the literature in any particular library. The first need is to measure the usage of the literature by those who contribute to it; the only publicly observable indicators of this usage are the references cited by the authors. The usage of a literature within any particular library will of course be of concern to the librarian, but his measures of usage will be of little interest to anyone else, nor are his data publicly observable. Citations within the literature to the literature, however, are publicly observable so that any results based on them can be checked by anyone who has access to the literature in any part of the world. If this publicly observable rate is known, then it should not be difficult to modify it to meet the usage patterns peculiar to any particular library.

The most straightforward technique of measuring obsolescence requires a sample of at least 2000 citations from the literature of the subject published within some specified year. The frequencies of citations to earlier papers *of the same literature* are then noted, taking the year as the unit. These frequencies are then cumulated in a table so that it is possible to read off how many citations were made *to each specified year or earlier*. The frequencies so cumulated, reducing year by year as the citations to still earlier years decline in frequency, are then plotted against their corresponding dates on semilogarithmic paper (two-cycle or three-cycle) with the frequencies marked along the logarithmic scale as in figure 5.

In general it will be found that after a hesitant start covering the two most recent years (because it takes that order of time to read a newly published paper critically, do the necessary scientific work, prepare the

paper and get it accepted and published), the plotted points fall steadily downwards on a straight line. This straight line continues until the period of the last world war is reached, when the plotted points usually move sideways (indicating an interruption of normal scientific publishing). Unless the sample is very large, the graph usually ends rather erratically as the frequencies become very small.

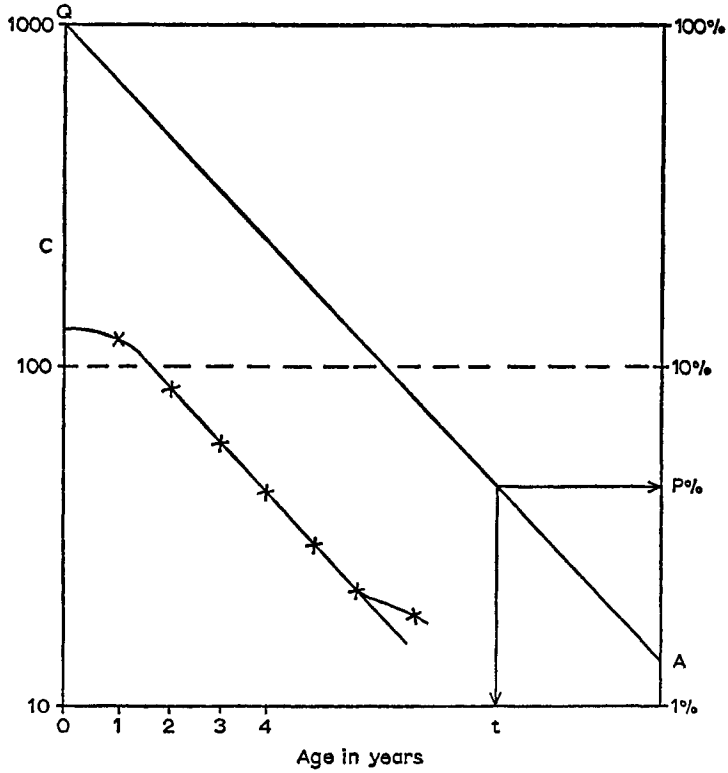

FIG. 5. Graphical Method of Estimating Obsolescence: C is the Number of Citations to the Year t or Earlier

The next step is to draw the best possible straight line through the *linear part* of the graph and to extend it to the edges of the graph paper. A second straight line QA is then drawn parallel to the first straight line and this second line becomes the means by which the annual aging rate and its effects can be read.

The theory of this procedure derives from the fact that the linearity

indicates that the citations as originally counted year by year can be expressed as the geometric sequence

$$R, Ra, Ra^2, Ra^3 \ldots Ra^{t-1} \ldots$$

where R is the presumed number of citations during the first year, some of which do not immediately emerge in publication. But as $a<1$, the sum of the sequence converges to the finite limit $R/(1-a)$. So the sequence plotted, which is the decreasing sum of citations to earlier years, can be expressed as

$$\begin{aligned} U_0 &= R(1 + a + a^2 + \ldots + a^{t-1} + \ldots) = R/(1-a) \ldots \qquad (6) \\ U_1 &= R(a + a^2 + a^3 + \ldots + a^{t-1} + \ldots) \\ &= Ra(1 + a + a^2 + \ldots + a^{t-2} + \ldots) \\ &= Ra/(1-a) = U_0 a \\ U_2 &= R(a^2 + a^3 + \ldots + a^{t-1} + \ldots) \\ &= Ra^2(1 + a + a^2 + \ldots + a^{5-3} + \ldots) \\ &= Ra^2/(1-a) = U_0 a^2 \end{aligned}$$

and so on. This result implies that the sequence plotted, i.e.,

$$U_0, U_0 a, U_0 a^2 \ldots U_0 a^{t-1} \ldots$$

is also a geometric series of the same ratio a as the original sequence. Hence, after t years the residual utility, $U_0 a^{t-1}$, is the fraction a^{t-1} of the original utility U_0. The value of this fraction can be read directly from the graph, either as a fraction or a percentage, as shown in figure 5.

Although the value of a can be read directly from the graph as the fraction corresponding to the age of one year, it is more accurate to read the fraction corresponding to the age eight years and then to take successive square roots. Thus if it is found that

$$a^8 = 0.36$$

then

$$a^4 = 0.60$$

and

$$a^2 = 0.775$$

giving

$$a = 0.88.$$

It is always helpful to draw the graph in case unexpected anomalies

arise which the librarian feels should be taken note of. Should any such anomaly make it difficult to draw any straight line with conviction, the constructed line OA cannot be drawn either. Any difficulty of calculation that arises, however, can be avoided by drawing a second graph beginning at Q which is geometrically parallel to the plotted graph. The required fractions or ages are then read by reference to the second graph. As the librarian can always exercise his judgment conservatively, there is no need for elaborate statistical procedures intended to yield figures of high precision.

If it has already been verified that, for the particular subject literature, the graphs can be assumed to be linear, it is possible to simplify the above procedure for estimating the numerical value of a. All that it is necessary to do is to divide the citation data into two categories when sampling and to count separately: (a) those eight years old or less and (b) those older than eight years. If these two counts amount to m and n respectively, then

$$m = R(1 + a + a^2 + \ldots + a^7)$$

and

$$\begin{aligned} n &= R(a^8 + a^9 + \ldots) \\ &= a^8R(1 + a + a^2 + \ldots). \end{aligned}$$

Hence

$$m + n = R(1 + a + a^2 + \ldots + a^7 + a^8 + a^9 + \ldots) = U_0$$

and

$$n = a^8U_0$$

so that

$$a^8 = \frac{n}{m + n}. \tag{7}$$

Square roots are then taken in succession as before to yield the required value of a.

It is emphasized that any value of a so obtained is the result of sampling and is therefore subject to sampling variance, i.e., two different samples of equal size drawn from the same literature are likely to give different results. Such differences decrease as the sample size increases but the sample size should always be stated when an obsolescence measure is determined.[7]

If the graph when plotted yields a slope which steadily decreases in

steepness, this result implies that the literature has at least two components of different rates of aging. A little knowledge of the subject usually makes it possible to separate the main components and to analyze them separately.

Obsolescence, Scatter and Growth

Librarians, concerned more with the usage of their own libraries than with the overall use of scientific literatures by those who create them, have been confused about the effects of the growth of literatures on their rates of obsolescence. They have argued that the value of a found by the technique described above has to be modified if the literature is growing. Their grounds for this argument rest on the idea that the probability of any particular paper being cited decreases as the number of papers published per annum increases. But they overlook the fact that scientific literatures grow, not because scientists increase their productivity, but because more scientists contribute to the literature at the same average rate. If this rate of growth of users of the literatures is taken into account, the probability of any new particular paper being cited remains constant. As far as earlier papers are concerned, the effect of growth is already implicitly included in the data collected. No corrections are needed.

In the usage of a *library,* however, other factors have to be considered. Thus, if, as in Britain, a new comprehensive library such as the NLLST is established, its usage will increase as its services become known and appreciated. It is possible that the rate of growth of usage of a scientific literature will equal or exceed the corresponding rate of obsolescence and the startling phenomenon of negative obsolescence may then be reported. It is, however, simpler to recognize that any literature ages at a uniform rate but that some libraries, especially new ones, can hope to attract usage at a rate which exceeds the rate of obsolescence. A simple analysis of the effect of growth on obsolescence is given in the Appendix.

There is some evidence, not yet wholly convincing, that scatter and obsolescence are positively correlated. It may be so, although at the present time there is no general agreement on how scatter should be defined or measured. The concept of scatter, except in the sense in which Bradford applied it to a comprehensive bibliography, is still vague. Yet there does seem to be a notion of scatter which is independent of completeness. For example, if 100 papers are selected at random from the literature of a subject and their sources are noted, one would expect that the number of sources so identified would be less than 100 and that the

fewer the number of sources, the less scattered the subject could be said to be. Unfortunately, the ratio of sources to number of papers selected at random can be shown to be a rather complicated function of the sample size. If the size of the sample of papers is increased from 100 to 200, the number of sources found in the sample of 200 is rather less than double the number found in the sample of 100. So, unless the sample size is standardized at, say, 100 papers, there is little prospect of finding a simple means of scatter by sampling. Although this problem is being studied, the related mathematics is surprisingly intractable.

It seems likely that scatter and obsolescence are related,[8] but that both are determined by rate of growth—the faster the rate of growth, the less is the scatter and the more rapid the obsolescence. A relationship between growth and scatter has been derived by Naranan[9] in mathematical terms, but bibliographic data are needed to test his theory in detail.

The literatures of most scientific subjects continue to grow exponentially, as judged by the number of papers published per annum, with a rate of growth over all subjects which has doubled the annual output in approximately ten years. The new literature makes its appearance dramatically in the copies of journals which a library receives every working day. The effects of obsolescence are much less dramatic: there is no external sign to indicate when the scientific utility of a journal has faded

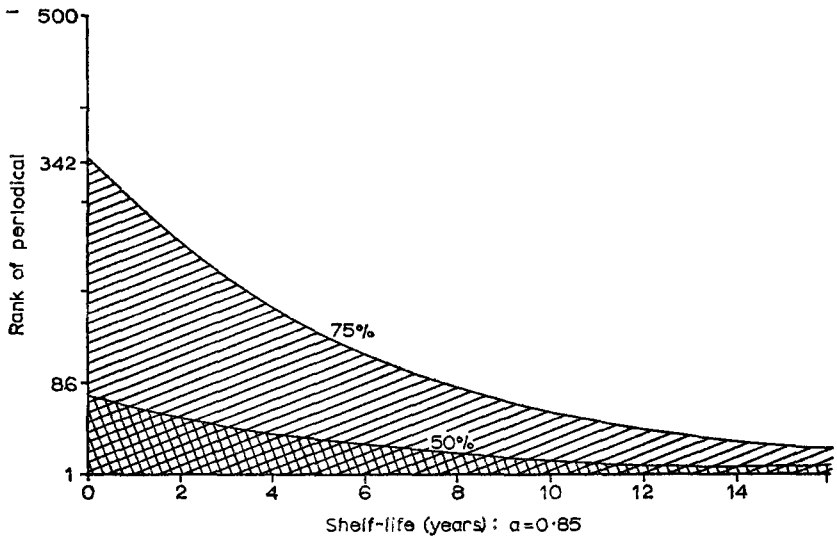

FIG. 6. The Contour Cut: Periodicals Discarded at Threshold Utility

into insignificance. Thus the literature of current scientific interest can be represented as the difference between an exponential growth and an equal exponential decay, but the growth line is well defined whereas the decay line is not. So the literature of *current interest* continues to grow but only as the difference between the curve of growth and the parallel but delayed curves of decay (figure 6).

The Optimum Specified Percentage of Utility for a Library of Specialist Periodicals

In this application Bradford's law is combined with the obsolescence law in the design of the most compact shelf-stock of periodicals which will provide any specified percentage of the utility obtainable from the acquisition and retention of the Bradford totality of N periodicals all with their complete book-runs.[10]

In the calculation that follows, it is assumed that all the relevant papers in the current issues of the complete set are of equal significance. This assumption is unlikely to be valid, for reasons mentioned earlier in this paper, but any error implicit in this assumption will yield a result which is on the conservative side.

If we now consider the utilities of the periodicals ranked as a Bradford set, we note that the productivity of the periodical ranked m (which lies outside the nuclear zone) is proportional to N/m. Thus the total utility of this periodical will be proportional to Nu/m. As all the periodicals age at the same rate, the more productive of two periodicals will always have the higher residual utility at any age t. If the library acquires the first m of the N ranked periodicals, then the periodical of rank m is at the minimum acceptable utility. The principle now applied is that the other periodicals are discarded when their residual utility declines to this minimum utility, i.e., equal residual utilities are discarded in the tails of each periodical.

For any other periodical of rank n, where $n \leq m$, we then note that it is discarded after t years where

$$Nua^t = Nu/m$$

which yields

$$a^t = n/m$$

so that the value of t can be read from the obsolescence graph.

The total utility lost by discarding the tails of the m periodicals will be $m\ Nu/m = uN$. The utility lost by not acquiring the periodicals of

rank $(n + 1)$ to N and discarding the periodical of rank m almost immediately will be

$$u(N \log N/s - N \log m/s) = uN \log N/m.$$

We can therefore write

$$uN + uN \log N/m = (1 - p)uN \log N/s$$

which expresses the fact that the utility discarded is equal to the fraction $(1 - p)$, where $p = P/100$, of the total utility of the complete Bradford set with all periodicals having full back runs. This equation yields

$$\log m/s = 1 + p \log N/s \tag{8}$$

from which m can be found for any specified value of p and known values of N and s.

Figure 7 shows the patterns of the $P\%$ library for $P = 50, 75$ when $N = 500$, $s = 2$, and $a = 0.85$. The nuclear periodicals are worth retaining without limit.

The "contour cut" just described is the most compact library, but it may not be the least costly because some of the periodicals of rank approaching m are discarded soon after acquisition. The simplest specified percentage of utility for a library is obtained by acquiring the requisite number of periodicals and discarding nothing. The appropriate number of journals in this case, m_1, is given by

$$N \log m_1/s = pN \log N/s$$

or

$$m_1 = N^p s^{1-p}. \tag{9}$$

Comparing (9) with (8) we see that $m_1 = m/e$ where e is the base of natural logarithms. So fewer periodicals are acquired but much more shelf space (and maintenance) is needed. The least costly solution is some compromise between these two solutions which depends on the relative costs of acquiring and of retaining and servicing the collection.

Although the contour cut solution may seem to be elaborate, it merely formalizes the way in which large collections of periodicals are dealt with practically in busy but efficient libraries where the users are allowed access to open shelves and where shelf space is limited. In such cases, the solution is determined by careful observation of usage rather than by measurement of obsolescence and application of Bradford's law. The volumes displaced from the open shelves are not discarded but are removed to closed-access stores nearby from which they can be re-

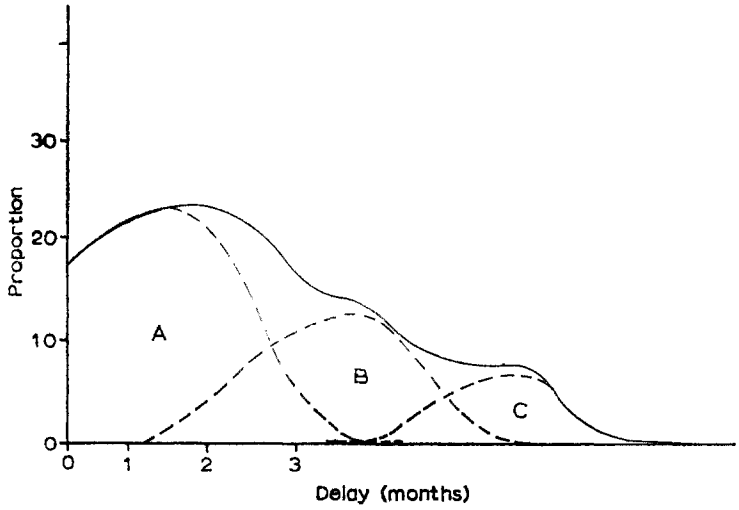

FIG. 7. Currency of SDI Services: A, National, B, Other, Same Language, C, Other, Different Language

trieved on request by members of the library staff. The problem here is therefore to insure that the volumes on open access are those most frequently sought by users. The principle applied is to remove to closed access those volumes whose average usage declines to some minimum frequency.

MEASURING THE EFFECTIVENESS OF RETROSPECTIVE SEARCHES AND CURRENT AWARENESS SERVICES

EFFECTIVENESS OF RETROSPECTIVE SEARCHES

Innumerable measures for quantifying the characteristics of information retrieval systems were proposed during the period of developing indexing languages and retrieval techniques during the 1950s and early 1960s. These measures were first applied to experimental manual systems with relatively small files and used with simple coordination of index terms which gave rise, for example, to Cleverdon's "inverse law of recall and precision."[11] Then Salton,[12] during his pioneering work on the mechanization of retrieval systems, introduced *normalized* recall and precision and the idea of ranking the items of the search output in terms of various measures of relevance. During this period, too, Swets[13] proposed theoretical models of information retrieval systems which were

again based essentially on the coordination of individual index terms drawn from a controlled list of terms.

Towards the end of that period the application of the computer to large data files became operational. And since that time, marked by the initiation of MEDLARS, the development of information systems has moved away from its obsession with the perfecting of index languages on sets of carefully prepared test-bed collections of well-indexed documents into an operational phase of development. Although indexing is still done mainly by technicians, the formulations of search questions have moved away from simple coordination of index terms into a more mechanical phase in which more sophisticated search languages, including syntactic components, such as the Boolean logical constants, have been developed. Step by step the computer has mechanized operations previously performed by human operators; first, the relatively easy matching of a group of terms representing the search question against groups of terms representing the documents; second, the elaboration of matching techniques, and now the replacement of the more intellectual processes of indexing by mechanized statistical techniques applied to titles and abstracts. The ultimate objective is the elimination of human operators, apart from the computer technicians, a task that will almost certainly be completed when input devices suitable for reading texts directly into the machine become more versatile, and indexing in terms of statistical analysis by machine of natural language texts becomes the norm.

In this phase of rapid technological development based on very large files and complex matching procedures, measures, empirical laws and theories derived from manual procedures and coordinate indexing have become increasingly inapplicable. The well-known average limits to effectiveness as measured by recall and precision seem to have become accepted as inescapably inherent in natural language descriptions. The only measure that remains directly applicable is *precision*, because that is the only measure that can be applied to the outputs of large mechanized operational systems. But precision by itself can be very misleading. Moreover there has been a shift of emphasis in operational systems from retrospective searching to current awareness services which are well suited to mechanization. As these powerful mechanized systems become established, new kinds of measure are needed; they will be more concerned with "user satisfaction" than with the efficiency of the searching mechanism itself.

The effectiveness of the retrospective searches of large mechanized

information systems is likely to be measured (unfortunately) in terms of their comprehensiveness. And, as a general test of the comprehensiveness of a bibliography, the application of the Bradford law, described earlier in this paper, is the only one yet available.

EFFECTIVENESS OF CURRENT AWARENESS SYSTEMS

Techniques of measuring user responses to current awareness (SDI) services are being developed by Leggate[14] and others. So far these techniques require the user to indicate which items he finds important, relevant but not important, and nonrelevant. Results of this kind can then be averaged and assessed using various conventional statistical techniques, although no single method has yet emerged which commands general acceptance. After all, the most critical test of operational systems run on a commercial basis is whether they can attract subscriptions from their users on a scale which makes them commercially viable.

One of the problems of concern to users of SDI services is that of currency. Users naturally expect the computerized SDI services they pay for to be up to date in their printouts. The user does not expect to see items which are a year or more old or which he knows have been superseded. The user of SDI services expects good currency rather than comprehensive coverage. Therefore, measures of currency are needed.

A practical difficulty in measuring currency is that the date of publication of a reference cited in an SDI printout, which is the nominal date of publication, may bear little relation to the date at which the reference was actually available in published form. So, in analyzing currency, the dates of the references should be measured, not from that quoted in the reference, but from the date of receipt of the source which, in an efficient library, is stamped on the source and corresponds to the day it was received. Random samples of the SDI output items can then be examined in relation to the date of availability and a frequency distribution of the delay in terms of months can be prepared. From such a statistical table two measures—the mean and the standard deviation—can be calculated, and both should be as small as possible.

Few data concerning currency have been published, but on the basis of the evidence available it seems that currency data take the form of truncated normal distributions as can most easily be shown by plotting the cumulative sums of delay times on arithmetic probability paper. If the SDI system offers coverage of worldwide sources it is interesting to separate the items which are yielded by the national sources of the country of the SDI service from those items which come from other

countries or from sources in other languages. In such cases one is liable to find two or even three overlapping normal distributions with increasing means and standard deviations (figure 7). Until the SDI system can accept direct inputs from the foreign sources, this aspect of currency is difficult to overcome; books and journals take appreciable time to be transported around the world, to be indexed or to be translated before abstracting and indexing.

One other irritant arises from the delayed publication of conference papers. These papers may bear the date of the conference but they may not be published in book or journal form for a year or more, and so they only then become available to the SDI service. Probably the solution to this difficulty arises from recognition of the different needs of retrospective searching and of current awareness. For comprehension the delayed conference papers need to be entered in the data base to be available for retrospective searches. But it is not necessary to compile the retrospective data base only by cumulating the periodic SDI collections which provide the current awareness services. The two services do not need to depend on precisely the same inputs.

Further Applications of Numerical Analysis

Goffman's general theory of information systems

The techniques of measurement applied to bibliography which have been described in the earlier sections of this paper have immediate practical application in the design of economic and efficient special libraries and bibliographical networks. But these measurements are also leading to quantitative descriptions of the characteristics of bibliographies which are gradually giving shape to what was regarded, only a few years ago, as an amorphous immeasurable confusion in which the seemingly limitless exponential growth of publication was the only quantified aspect. These quantitative descriptions are needed for the development of theories about scientific communication. If information science is to become the science it explicitly announces itself to be, it needs to develop a theory which lifts the subject off its present ground of ad hoc technological development and gives it autonomy, depth and more creative objectives.

A new area of analytical interest has been opened up by Goffman,[15] who has been exploring in detail the analogy between the dissemination of scientific knowledge and the spread of epidemics. He has analyzed the complete bibliographies of several emergent scientific descriptions

over periods of 100-200 years. He has shown that some of the ideas were "endemic" with minor outbreaks occurring from time to time. A man with some new knowledge is rather like the carrier of an infectious disease, but the disease is transmissible only to "susceptibles" and there must be a critical density of susceptibles for an epidemic to occur. The importance of this work lies in three directions: (1) it provides a new empirical analysis of the growth of science which should illuminate the recent philosophical descriptions of this process; (2) it offers a means of predicting whether an epidemic growth of an emergent subject is likely to occur or not; and (3) it establishes a basis for a much-needed general theory of communication which should give direction to the development of the worldwide information systems that are now being rather superficially discussed.

The theory Goffman is creating rests on two established theories—on epidemiological theory and on the Shannon information theory[16]—which are applied together in the analysis of the information phenomena recorded as comprehensive scientific bibliographies. The understanding of this theory requires competence in mathematics and statistics and cannot be summarized within the scope of this present article. It is mentioned here because it is already providing spin-offs of immediate practical interest. One of Goffman's needs is the compilation of comprehensive scientific bibliographies which in itself is an exercise of practical interest since it reminds us that the existing computerized data banks can offer little help with this project and that there is a need for a world repository of the tapes on which these hard-won bibliographies are now being recorded in FAMULUS format. Such tapes, suitably tagged for all foreseeable kinds of bibliographical analysis, are needed to increase the number of susceptibles that Goffman's own epidemiological theory of scientific growth requires for the epidemic growth of information science.

THE COMPACTION OF BIBLIOGRAPHIES

Although the computer has come to the aid of the bibliographer as a mechanism for compiling, editing, storing, and sorting the items that constitute the bibliography, it has done nothing to help in the task (which is not strictly that of the bibliographer) of sifting the grain from the chaff. The present ideal of the computerized data base is also that of the bibliographer—to compile *everything*. But if that means that the scientist in hot pursuit of some scientific objective is always going to be supplied with *everything* that is "relevant" whenever he asks for a retro-

spective search, he is going to be stopped in his tracks. The growth of science is all too often regarded as the steady accumulation of all knowledge, a view which is symbolized by the ever-growing data banks of the operational computerized systems. But. in fact, science is the continuous reappraisal of the old in terms of the new which discards the old almost as fast as it generates the new.

Until computerized data banks can reflect this more realistic view, their scientific users will become increasingly disillusioned with the lengthening printouts they receive. Work is therefore underway to develop methods of compiling, filtering and compacting scientific bibliographies so that current information can be presented in an immediately useful assimilable form. Information on tape ages at the same rate as the same information stored in archival collections of printed periodicals. No useful purpose is served by the continued revival of obsolete material, even when it is done by computer.

APPENDIX

The Relation Between Growth and Obsolescence

It is known that a geometric sequence is obtained from the citation dates obtained from a homogeneous literature whether the literature is growing or not. So, in general,

$$U_0 = R_0/(1 + a + a^2 + \ldots + a^{t-1} + \ldots) = R_0/(1 - a). \tag{1}$$

If the literature is growing exponentially at the rate g per annum, then after time t, R_t is given by

$$R_t = R_0 e^{gt}. \tag{2}$$

If the number of contributors (i.e., users) is also growing exponentially but at rate s per annum, then

$$U_t = U_0 e^{st}. \tag{3}$$

Differentiating (1), (2) and (3) with respect to t, substituting from (2) and (3) and simplifying, we have

$$\frac{da}{dt} = (s - g)(1 - a). \tag{4}$$

Clearly, if $s = g$, the value of a, as measured by citation counts, should remain constant. If $s \neq g$, then, on integrating (4), we have

$$(1 - a_t) = (1 - a_0)e^{(g-s)t} \tag{5}$$

where a_0 is the value of a when $t = 0$.

This result implies that while $g \neq s$ the value of a_t will change from year to year.

In the only empirical test of (5) yet published, Oliver[17] found that s and g *were* equal (within the limits of sampling error) for the literature of solid

state physics and that $a = 0.79$ at both the beginning and the end of a five-year period of rapid growth.

References

1. Bradford, Samuel C. *Documentation.* London, Crosby Lockwood, 1948.
2. Wilkinson, Elizabeth A. "The Ambiguity of Bradford's Law," *Journal of Documentation,* 28:122-30, June 1972.
3. Brookes, B. C. "Bradford's Law and the Bibliography of Science," *Nature,* 224:953-56, Dec. 6, 1969.
4. Goffman, William, and Morris, Thomas G. "Bradford's Law and Library Acquisitions," *Nature,* 226:922-23, June 6, 1970.
5. Brookes, B. C. "Photocopies v. Periodicals: Cost-effectiveness in the Special Library," *Journal of Documentation,* 26:22-29, March 1970.
6. ———. "The Design of Cost-effective Hierarchical Information Systems," *Information Storage and Retrieval,* 6:127-36, June 1970.
7. ———. "Obsolescence of Special Library Periodicals: Sampling Errors and Utility Contours," *Journal of the American Society for Information Science,* 21: 320-29, Sept. 1970.
8. Buckland, Michael K. "Are Obsolescence and Scattering Related?" *Journal of Documentation,* 28:242-45, Sept. 1972.
9. Naranan, S. "Power Law Relations in Science Bibliography: A Self-consistent Interpretation," *Journal of Documentation,* 27:83-97, June 1971.
10. Brookes, B. C. "Optimum P% Library of Scientific Periodicals," *Nature,* 232: 458-61, Aug. 13, 1971.
11. Cleverdon, Cyril W. "The Cranfield Tests on Index Language Devices," *Aslib Proceedings,* 19:173-94, June 1967.
12. Salton, Gerard. *Automatic Information Organization and Retrieval.* New York, McGraw-Hill, 1968.
13. Swets, John A. "Information Retrieval Systems," *Science,* 141:245-50, July 19, 1963.
14. Leggate, P. *Evaluation of Operational Current Awareness Services* (OSTI Report 5111). 1971. (Deposited in NLLST)
15. Goffman, William. "A General Theory of Communication." *In* Tefko Saracevic, ed. *Introduction to Information Science.* New York, R. R. Bowker, 1970.
16. Shannon, C. E. "The Mathematical Theory of Communication," *Bell Telephone System Journal,* 27:379-423, 623-58, 1948.
17. Oliver, Merrill R. "The Effect of Growth on the Obsolescence of Semiconductor Physics Literature," *Journal of Documentation,* 27:11-17, March 1971.

The Humanities—A State of the Art Report

LAWRENCE S. THOMPSON

VOLUME, conditions of production, distribution, use, and bibliographical management of publications in all categories have become vastly more complicated in the twentieth century than ever before. Proliferation of research institutes and vastly increased financial support in all fields (but, last of all, the humanities) are striking phenomena of scholarship in our time. Problems this situation has created are both relieved and aggravated by the rapid development of mechanical technology such as microreproduction, sophisticated devices for composition and offset printing, and new systems for indexing and for compiling bibliographies. All are being constantly refined; in turn, the geometrically increasing volume of subpublication and the creation of new indexes generate new research and publication. Here is an apparently unending cycle, and the critical problem is bibliographical management of masses of material which might otherwise be lost, or at least become most difficult of access.

Methods of bibliographical control of the humanities did not change much between Konrad Gesner and Joseph Sabin. In the past monumental bibliographies have generally been the work of one man, e.g., Gesner, Hain, Brunet, Sabin, and Evans, although frequently their work must be continued by others (Sabin) or supplemented (Hain). In our time there are still the individual bibliographers, and they will, hopefully, always be with us; but there is a growing trend for major bibliographies to become "projects," undertaken only by teams and then only with the prerequisite of substantial funds and often elaborate technical equipment. This arrangement is just as well whenever it is practical, for it relieves the creative, imaginative bibliographer of chores of which he is not worthy, or releases him for those for which only he is competent.

Detailed, individual bibliographical scholarship is here to stay unless

Lawrence S. Thompson is Professor, Department of Classics, University of Kentucky, Lexington.

the whole definition of research in the humanities is changed. It would be as ridiculous for a textual scholar not to use a Hinman collator as for an administrative office not to use an offset machine; but the individual scholar's direct control over the whole process is essential. One egregious example (of many) is Matthew J. Bruccoli's *F. Scott Fitzgerald, a Descriptive Bibliography* (1972) requiring the concentrated effort of a single dedicated individual who is at once a scholar, a bibliographer, a collector, and a bird dog for minutiae buried in all manner of unlikely places. Indeed, monumental works such as the *Gesamtkatalog der Wiegendrucke* and the *Catalogue of Books Printed in the XVth Century Now in the British Museum* are largely the work of individuals pulled together by a single agency or library, a fact clear to any reader of bibliographical journals. The pages of serials such as *The Library* and the *Gutenberg Jahrbuch* are strewn with the timber from which the great catalogs of incunabula were built.

Always a special problem of humanistic and historical scholarship has been to identify quality or lack of it; and in no area is it so easy for the careless scholar to escape detection (for a while), or so easy for the superior scholar to be denied recognition he deserves in his time. One solution is to include everything in sight and let the reader make his own qualitative evaluations, the case with Hensley C. Woodbridge, *et al., Jack London: A Bibliography* (1966), a work which even runs down often insignificant reviews and partial translations into obscure languages. Yet here is the material from which we can explain London's inflated reputation outside of North America. The other extreme involves a sovereign mastery of the field involved, identification only of cornerstone works such as we find in the late J. Christian Bay's masterful "Three Handfuls of Western Books." No mechanical device can make the decisions that Woodbridge or Bay faced.

A factor of increasing importance for humanistic bibliography is the official and private subsidy. Only the Lilly Foundation has made the *Bibliography of American Literature* possible; and many another bibliographical work acknowledges the assistance of foundations, great and small (not all of which spend their money as wisely as Lilly). UNESCO has aided very substantially in compilation and publication of bibliographies of translations, philosophy, history of religian, color reproductions of paintings, and other subjects; and, on a national level, the Deutsche Forschungsgemeinschaft, the Centre National de la Recherche Scientifique, and the National Endowment for the Humanities (much more poorly supported, relatively speaking) have had much the

same role. In general, however, funds are likely to go to individuals and groups which are most familiar with the routines, formalities, and verbal expression for making requests, not necessarily to the most urgently needed works. In the age of the welfare state there is a new meaning for Ovid's "inopem me copia fecit."

An element which always has and always will plague bibliographers is the matter of currency. Every bibliography is out of date the moment it stops at a *terminus ante quem.* The impatience of scholars for a bibliographical record often creates duplication through the compilation of "preliminary handlists" and splinter bibliographies for specialists in a narrow aspect of a larger field. Fortunately, current bibliography is much more common in this century than ever before. It was not sufficient for W. F. Poole to index nineteenth century periodicals, and it was necessary to set up the *Readers' Guide to Periodical Literature* on a regular, current basis. In a subject field such as classical studies it was not satisfactory to have at hand only the sequence of noble bibliographies by Engelmann, Klussmann, Lambrino, and Marouzeau. *L'Année philologique* had to be created to assure scholars of comprehensive coverage of their field on a current basis. Most other fields of the humanities have also come under current bibliographical coverage in the twentieth century, although some, such as linguistics, are relatively late, e.g., the *Bibliographie linguistique,* beginning publication only in 1949 but, happily, covering 1939 to date. It seems unlikely and is probably not even desirable for any field in the humanities to aspire to immediate bibliographical coverage, imitating, for example, the bibliographers of the medical trades, since the quality of work in the humanities varies so widely.

We have noted that the eagerness of twentieth-century scholars for current bibliographical coverage of their fields and up-to-date reference books often results in duplication and overlap, but in many cases this situation is justifiable. Selective bibliographies for students and nonspecialists are necessary to help the inexperienced and to provide a point of departure for the old hands. Thus the classical scholar who needs to have a quick conspectus of some area with which he is not familiar need not wade through Pauly-Wissowa and *Der kleine Pauly* (1964-date), for he can generally find immediately needed references in the *Lexikon der alten Welt* (1965) or the second edition of the *Oxford Classical Dictionary* (1970). Several decades ago the British Museum found that it was necessary to issue short-title catalogs of its pre-1601 imprints in various jurisdictions, a duplication of its general cata-

log and of its catalog of fifteenth-century imprints. Today this activity is being extended with V. F. Goldsmith's *A Short-Title Catalogue of French Books, 1601-1700, in the Library of the British Museum* (1969-date). On the other hand, Suzanne and Paul-Henri Michel's *Répertoire des ouvrages imprimés en langue italienne au XVII siècle conservés dans les bibliothèques de France* (1967-date) includes many titles and copies not in the Bibliothèque Nationale and recorded in its catalog. Hopefully Goldsmith and similar works can be expanded to or supplemented by a similar union catalog and show locations even beyond those in the book form *National Union Catalog* (currently being printed at a fantastic but justifiable price by Mansell of London).

A more reprehensible source of duplication is generated by scholars who publish books and accompanying bibliographies on the same persons or subjects. At times such duplication is proper when one study is patently inferior; but too many scholars are too stubborn to give up a research project even when they learn at an early stage that there is a competitor. Even a brief check of the bibliographies of bibliographies in various jurisdictions, language areas, and subject fields will show how widespread this situation is; but there is little that can or, indeed, should be done about it.

In the remaining parts of this article it will be possible to provide only a resumé of current trends in selected fields of humanistic bibliography and to indicate how they reflect some of the conditions noted in the earlier paragraphs. Older works will be noted only insofar as they are pertinent for the later twentieth century. Two previous issues of *Library Trends* have been put together as *Bibliography, Current State and Future Trends* (1967), edited by Robert B. Downs and Frances B. Jenkins, and provide much more detail than can be given here. Examples selected for discussion will be from genres of bibliography (e.g., dissertation and periodical indexes, bibliographies of homage volumes, comprehensive and partial library catalogs), technology and its effects on bibliographies (e.g., microforms and catalogs thereof, cheap reprinting in eye-readable form, and computers and related instruments), trends in humanistic scholarship (e.g., the newer textual criticism), and a few selected fields such as literary history, linguistics, philosophy and religion, and music which have been or may be affected by conditions (old and new) for humanistic bibliography and by recently developed methods and aids to scholarship.

A special plague of librarians and bibliographers from the sixteenth century to the present has been the academic dissertation. The problem

has been more aggravated in the humanities than in the sciences, since the results of doctoral research in the latter usually appear as journal articles, whereas a large proportion of those in the humanities are buried in manuscript form in the libraries of institutions where they were presented. There are, to be sure, national lists such as the *Jahresverzeichnis der deutschen Hochschulschriften* (1885-date) and, much more recently in other major countries such as the United States (1933-date), Great Britain (1953-date), and the U.S.S.R (1958-date). However until quite recently, dissertations as a genre have not been analyzed in detail by subject. During the last decade there have been several useful bibliographies which meet the needs of the modern subject specialist, although most of them have been confined to the United States, Canada or Germany. Such is not the case with one of the most recent, Leonard H. D. Gordon and Frank J. Schulman's *Doctoral Dissertations on China; A Bibliography of Studies in Western Languages, 1945-1970* (1972), the first number in a series initiated by the Association for Asian Studies. Hopefully there will be complementary volumes on pre-1945 titles and Chinese and Japanese titles. Other occidental lists would do well to include oriental, above all, Japanese doctoral work. Although works such as Gordon and Schulman pull together material that is otherwise recorded only in institutional or national lists, neither it nor most of the others are indexed in depth. Such is not the case with the present writer's *Bibliography of American Doctoral Dissertations in Classical Studies and Related Fields* (1968) in which there are four subject entries to each author entry. It was suggested to the compiler that a system such as KWIC be used for indexing, but it was soon ascertained that this device was only a quick way to avoid the scholarly bibliographer's primary responsibility for adequate subject indexing. To visualize the KWIC system's treatment of a title such as *The Elder Turtles of Aigina* (on numismatics) would be shock treatment for the imagination. Mechanical indexing has a rocky road ahead in most aspects of humanistic subject bibliography, but there is an undeniable potential.

A somewhat similar problem is posed by homage volumes, especially when they are not special numbers of periodicals or other serials indexed regularly in the appropriate organs. The older Festschriften provide a special challenge, and it has been met in several fields in recent years. Dorothy Rounds did the job for classicists with her *Articles on Antiquity in Festschriften* (1962), and Harry F. Williams did it for a subsequent period with *An Index of Mediaeval Studies Published in*

Festschriften, 1865-1946, with Special Reference to Romanic Material (1951). Herbert H. Golden and Seymour Simches did the same for French, Italian, and Iberian literatures and languages in a series of bibliographies published in the 1950s. Students of English, German, Scandinavian, Slavic, and Celtic studies need similar tools. Where, for example, can one find references to the treasury of studies on mid-American speech in Gordon Wilson, Sr., *Folklore of the Mammoth Cave Region* (1968), edited by Lawrence S. Thompson for the honoree's eightieth birthday?

Periodical indexing in the humanities has never been as satisfactory as it is in the sciences. Even a massive work such as *Chemical Abstracts* is reasonably sure to cover the great majority of the significant literature in the field, but no index in any field of the humanities can boast of the same adequacy. The sciences are measurable, subject to definition in precise terms, the humanities are not. Further, most serious journals in the physical and biological sciences are refereed, a policy adopted widely in the humanities only in recent years. There are literally thousands of serials which print articles, reviews, and notes at the whim of the editor. They are good, bad, and indifferent, mostly of the latter two categories. But how are they to be evaluated for purposes of the bibliographical record? The corpus is so vast that only students in specialized fields such as classical epigraphy or prototypography (both of which have their own bibliographies) can easily identify that which is essential and scholarly. In a field such as national literary history the way of the bibliographer is even more deceptive: a serious study of a regional author may appear in the organ of a local historical society, or a critical analysis of a composer of sectarian hymns could be encaved in the pages of the bulletin of a little-known theological seminary.

There will always be a place for the *Readers' Guide to Periodical Literature* in public libraries, and scholars in any field of the humanities will neglect it at their own peril. So too we need indexes of a national scope, particularly for smaller countries, such as the *Dansk Tidsskrift-Index* and the *Index to South African Periodicals*. Whether the monumental *Internationale Bibliographie der Zeitschriftenliteratur* justifies the expense of compilation and the cost to libraries for acquiring, shelving, and interpreting it to confused readers is questionable. A large proportion of the material is indexed elsewhere on a current basis, and specialists can find it. The general reader is easily satisfied with indexes of much more modest proportions and reviewing organs which are available in the average research library. Comprehensive indexing

of all material in serials, with adequate subject analysis, is a desideratum with a fairly low priority. It is not feasible with human and technical resources available at present, but it is not beyond possibility. An alternative, not beyond possibility either, is authoritarian suppression of unimportant serials.

The demise of the old *Bibliographie der Rezensionen und Referate* (1900-1943) was a blow of major consequence to humanistic scholarship. Reviews, however uncritical, brief, or apparently insignificant, are a fundamental key to the reputation of an author or the spread of an idea; and the critical reviews frequently add substantially to the subject matter at hand. The *Book Review Digest* (1905-date) and the *Letopis' retsenzii* (1935-date) are national in scope; and the national periodical indexes and subject bibliographies are not consistent in listing reviews. There seems to be no practical solution at present for providing comprehensive access to reviews in the humanities.

The last quarter of a century has witnessed an enormous advance in the production of library catalogs, the fundamental bibliographical tools for students in the humanities and social studies. When the 167 volumes of *A Catalog of Books Represented by Library of Congress Printed Cards* appeared in 1942-46, it was not only a great accomplishment in itself but also a challenge to get other basic catalogs (above all, that of the British Museum) into print. G. K. Hall of Boston and Mansell of London have performed important services for humanistic scholarship in this field. Catalogs such as those of the Hispanic Society or the Wing Collection in the Newberry Library provide bibliographical information and much more direct access to these great collections than scholars have ever had before. The *Deutscher Gesamtkatalog* was perhaps the major bibliographical casualty of the 1939-1945 war; and, while there is little hope of its resumption (showing, *inter alia*, books recorded in the first fourteen volumes which no longer exist), the very useful German regional union catalogs might someday be interfiled and printed in book form. Perhaps more immediately practical are works such as the *Svodnyi katalog russkoi knigi grazhdanskoi pechati XVIII veka* (1962-1966), a well-nigh complete catalog of eighteenth-century Russian imprints. Eighteenth-century scholars in western Europe are urgently in need of similar works for their linguistic areas. Such a project has been widely discussed for several years by students of eighteenth-century English literature.

A category of library catalogs of particular interest to humanistic scholars are catalogs of incunabula. Frederick Goff's *Incunabula in*

American Libraries; A Third Census . . . (1964) is the third successful effort to record fifteenth-century imprints in North American libraries; and the strong backing of the sponsor, the Bibliographical Society of America, makes a fourth census a reasonable certainty in due time. On the other hand, the *Gesamtkatalog der Wiegendrucke* presents problems of considerably greater magnitude. Halted at the end of Lieferung 1 of volume VIII, the *GW* is now underway again with the valiant efforts of the staff of the Inkunabelabteilung of the Deutsche Staatsbibliothek (East German torso of the old Preussische Staatsbibliothek). It receives relatively little support other than from libraries, individual scholars, and the publishers, Hiersemann of Stuttgart and H. P. Kraus of New York. The great foundations and the official agencies of all nations have invested far greater sums in projects of far less value.

A serious problem for humanistic scholarship is bibliographical control of microforms. When microforms became widely used in the late 1930s, they were produced in geometrically increasing quantities, with little thought of the problems of identification and cataloging. One major project, that of the microfilm edition of English books printed before 1640, based on A. W. Pollard and G. R. Redgrave's *A Short-Title Catalogue of Books Printed in England, Scotland and Ireland and of Books in English Printed Abroad to the Year 1640* (1926), was accompanied by catalog cards, but these cards were produced on the initiative of and by an institution, not the publisher. Fortunately, the film is identifiable by STC number and can be retrieved simply by checking a copy of the STC; but this is not true of many other microform projects of staggering proportions. One, with well over 10,000 titles, even attempts to remedy the deficiency of cataloging by printing a form card with the name of each author in the collection, indicating that all of his works published in the chronological and regional scope of the project are available in the collection! Today the gospel of providing catalog cards to accompany microform editions has still not yet penetrated to all publishers. Alone among microform publishers at this time, the Erasmus Press, Falls City Microforms, the Lost Cause Press, and General Microfilm provide catalog cards for all of their publications, without exception, running to some 8,000 titles in 1972.

In addition to providing catalog cards, each group should be accompanied by a list, and these lists should be cumulated, in full bibliographical detail, into book catalog with supplements as necessary. The first catalog of this type was that of G. William Bergquist, *Three Centuries of English and American Plays* (1963), recording some 5,000 ti-

tles in the project to copy on "microprint" British plays, 1500-1810, and United States plays, 1714-1830. The French, Spanish, and German plays issued by Falls City Microforms, much larger in scope in each instance, are represented by printed catalogs, and supplements are in preparation. Falls City's catalog of its English and American plays of the nineteenth century on microfiche is in press, and so too is its catalog of French revolutionary pamphlets on microfiche. A very important by-product of adequate catalog cards for microforms is that the way is then easy to inclusion of all titles in the *National Union Catalog of Microforms*. Bibliographical control of microforms on all fronts is a critical problem for libraries today; and, as one frustrated scholar faced by thousands of improperly indexed and cataloged rolls said, he would prefer the chore of *mulgere hercum*.

Along with microforms, the major technical development of our times is mechanized information retrieval. Although many devices such as edge-notched cards, see-through or "feature" cards, and related systems can often be applied to problems in humanistic research, we can discuss here, and then only very briefly, the impact of the high-speed digital computer and its offspring and elaboration in the last two decades. Its impact on the field of humanism has been mainly in the area of indexing and compilation of concordances, although it also has a potential use for detailed textual studies. Matters such as variant texts and readings, compositors' errors of all varieties, or all types of typographical problems might well be subjected to computer analysis; but only the human machine can identify ghosts, deliberate interpolations or deletions, or psychological conditions of the author and compositor.

The journal *Computers and the Humanities*, published at Queens College of the City University of New York in Flushing, has some highly suggestive articles. For example, a study of a project for a concordance to *Faust* (IV [1970], 161-171) by Theodore C. Hines, Jessica L. Harris, and Charlotte L. Levy offers a method which can readily be applied to other similar literary works. The series of *Deutsche Wortindices* being issued by de Gruyter in Berlin has covered Büchner's works and Keller's *Die Leute von Seldwyla* and *Der grüne Heinrich* thus far. The machines have produced work that a Cruden or a Bartlett might have relegated to underpaid amanuenses, but with considerably greater accuracy. The final printout is not as pretty as a typographical job, but it is equally practical. On the other hand, no aspiring young scholar should view the editing of a computer-produced concordance as a quick road to fame. The editor must apply the same critical judg-

ment to the preparation and execution of the work as did Cruden or Bartlett. But he will not be responsible for routine chores of which the mind of man is unworthy, and can apply himself to greater refinement of the job at hand. Such has been the role of mechanical devices in the world of the intellect ever since the chisel was supplanted by the calamus.

The flowering and refinement of two older technologies, offset printing and photography, is the basis of the modern eye-legible reprint industry; but the great expansion of the reprint business in the last two decades has been due mainly to the proliferation of higher institutions and their libraries in Western Europe and America and the increased financial resources of the older ones. Again here, scientific works were the first to be reprinted; but in the 1960s there was a veritable deluge of reprints in the humanities and social studies, estimated to be equal to all previous reprints since Senefelder's invention was used by book publishers. In particular humanistic bibliography has benefited, and not simply by having back in print fairly recent books. Thus the 1755 edition of Lione Allacci's *Drammaturgia,* never superseded, tedious as it may be to use, has been reprinted by the Bottega d'Erasmo of Torino. Several of the invaluable bibliographies of nineteenth-century authors by Thomas James Wise are now available again, a situation about to trigger reprints of the Wise forgeries themselves! Medina's invaluable Hispanic-American bibliographies, nearly all on wretched paper, are now available in much more durable form. In general the reprinting of humanistic bibliographies has been more judicious than reprinting in most other fields, although there have been a few bibliographies that might better have been put aside for revision and expansion.

A significant trend in literary scholarship which is having a profound effect on descriptive bibliography are the methods of textual analysis developed by Walter Greg and Fredson Bowers and his school. For a half century now bibliographical scholars have been examining the texts of older English writers and bringing to bear every technique of typographical and other physical evidence, as well as circumstances of composition and transmission of the author's original, to establish reliable texts. There has been an especially strong emphasis on this variety of bibliography in the United States; and it has spread wherever Bowers's students have moved from its *fons et origo* at the University of Virginia. For example, a group of bibliographical scholars at the University of South Carolina has established a counterpart to *Studies in Bibliography* (organ of the Bibliographical Society of the University of

Virginia) in the form of *Proof, the Yearbook of American Bibliographical and Textual Studies* (1971-date). The "Pittsburgh Series in Bibliography" from the University of Pittsburgh Press has already produced noteworthy bibliographies of Hart Crane and F. Scott Fitzgerald in this new tradition. The Indiana edition of Howells and the Ohio State edition of Hawthorne have applied the Greg-Bowers methods to classic American authors with salubrious results as far as the establishment of a proper text is concerned. In the future historians and critics of literature will be treading on dangerous ground when they do not have reliable texts at their disposal.

The enumerative bibliography of current literary scholarship has been revised and reorganized substantially in the last quarter of a century. At the end of the 1930s only classical and German studies, among literary disciplines, had adequate current bibliographies in *L'Année philologique* and the *Jahresbericht über die wissenschaftlichen Erscheinungen auf dem Gebiete der neueren deutschen Literatur* (1921-36/39). The expansion of the old *MLA Bibliography* to the *MLA International Bibliography* in the years after 1955 has been the most significant event of our time in the field of enumerative literary bibliography. It indexes well over a thousand serials as well as books, Festschriften, and other vehicles of scholarly publication. It is not difficult to find secondary or ephemeral work cited here, but what editor or editors can easily develop a better policy?

As we have noted before, the specialist is rarely satisfied by omnibus bibliographies covering all of his broad field on account of their tardiness, their lack of critical perspective in many cases, and their omissions. The Renaissance scholar, for example, has been nourished from his academic cradle on the annual bibliography of his field in *Studies in Philology* and is not eager to have to excavate his references from a massive work fifty times its size. The annual *American Literary Scholarship* (1963-date) states in the foreword that there is active collaboration with the *MLA International Bibliography*'s editors, thus frankly admitting the need for some little duplication. On the other hand, *American Literary Scholarship* continues a long tradition of critical, narrative bibliography begun a century ago by Bursian's *Jahresbericht* (continued by *Lustrum*, 1957-date) for classical studies, the *Year's Work in English Studies* (1919/20-date), the *Year's Work in Modern Language Studies* (1929/30-date), and *Germanistik* (1960-date). Entries may be duplicated, but the critical analysis is an urgent need of serious scholars.

Much more distressing is the duplication of the biennial *Bibliographie der deutschen Literaturwissenschaft* (Frankfurt am Main, 1945/53-date) and the *Jahresbericht für deutsche Sprache und Literatur* (Leipzig, 1940/45-date). However, there are other equally grave duplications of bibliographies in the two Germanies, notably the national bibliography, and it must be recognized (with regret) that the bibliographer cannot influence the politician.

One of the major improvements in coverage of current publications on a national literature is Otto Klapp's *Bibliographie der französischen Literaturwissenschaft* (1956/58-date). There was, of course, the useful bibliography in the *Revue de l'histoire littéraire de la France* (1894-date), but it was far from being as extensive in coverage as the work established by Klapp. Presumably the *Repertorio bibliografico della litteratura italiana* (2 vols., covering 1948-1953) will be continued, just as it continues Prezzolini; but a current bibliography is needed urgently. Hispanic studies are in a much worse shape. If the old *Bibliographie hispanique,* issued annually from 1905 to 1917 by the Hispanic Society, could be revived on a broader basis to include all Hispanic and Luso-American areas, a serious lacuna in modern literary bibliography would be filled. The splintering of Hispanic culture and the lack of a recognized metropolis is a deterrent. In the Soviet Union *Novaia sovetskaia literatura po literaturovedeniiu* and *Novaia inostrannaia literatura po literaturovedeniiu* not only cover Russian literature adequately but also pick up a great deal that is missed by the MLA.

Linguistic bibliography was largely confined to works for students of single language or groups of languages until the *Bibliographie linguistique* (1948-date) began to appear annually in 1951, supplementing the two volumes of the *Bibliographie linguistique des années 1939-1947* (1949-1950). It covers the languages of the world, thus overlapping to some degree the *MLA International Bibliography* and bibliographies of national literatures and languages. Most appropriately, it has been supported by UNESCO. From 1913 to 1948 there was adequate critical bibliography of Indo-European languages in the *Indogermanisches Jahrbuch,* a serial continued by *Kratylos; kritisches Berichts- und Rezensionsorgan für indogermanische und allgemeine Sprachwissenshaft* (1956-date). Linguistics, perhaps as much as any discipline, needs a current critical review of the literature, particularly in view of the uncertainty and divergence among scholars about such matters as the language of Linear B or the languages of Asia Minor in the second millennium B.C. Further, linguistic bibliography needs to expand to coverage

of related literature in biological and behavioral sciences, areas which have not always had the full respect of linguists in the older tradition.

Philosophy is another field which has enjoyed the budgetary favors of UNESCO. The useful *Répertoire bibliographique de la philosophie* (1949-date), published quarterly in Louvain by the Institut Supérieur de Philosophie indexes virtually all important philosophical journals and also picks up pertinent articles in other journals. There is a list of book reviews in the November issue of each year. The *Bibliographie de la philosophie* (1937-1958; superseded by a quarterly of the same title, 1954-date) is now an abstract journal covering books only. Section 19 of the *Bulletin signalétique* of the Centre National de la Recherche Scientifique (C.N.R.S.) covers philosophy and indexes journals from the entire world, giving brief abstracts in French. Despite some duplication, the current bibliography of philosophy is in good condition, certainly much better than it was a half century back when the philosopher had to depend on Ueberweg and several current sources of bibliographical information.

UNESCO has also favored a related field by subsidizing the *International Bibliography of the History of Religions* (1954-date), issued under the auspices of the International Council for Philosophy and Humanistic Studies and the International Association for History of Religions. It records journals and books and includes all major religions, ancient and modern. Unfortunately, it is quite tardy in appearance, a factor which limits its usefulness as a current tool. Since 1949 the American Theological Library Association has been issuing the *Index to Religious Periodical Literature*, a work which has grown considerably in scope but is still far from providing adequate coverage of the field. Retrospective bibliography of the history of religion and of theology in general must be picked up from a variety of sources; but the best points of departure are the third edition of *Die Religion in Geschichte und Gegenwart* (1956-1962) and the second edition of the *Lexikon für Theologie und Kirche* (1957-1966). There is considerable overlap in the bibliographies of philosophy and religion, but it would be difficult to avoid. Scholars in both fields need their own bibliographies, current and retrospective.

The bibliography of music and musicology is rich, although it faces several problems, especially the matter of being up to date. Both the *Bibliographie des Musikschrifttums* (1936-date) and the *Music Index* (1949-date) are quite late in appearance. The latter indexes some 200 periodicals, but the former also includes books, reviews, and essays in

sources other than journals. A subsidy from some official source might help both to be more prompt. Retrospective music bibliography must be gathered from various sources, but *Die Musik in Geschichte und Gegenwart* (1951-1968) will get the scholar off the ground in most subjects. The bibliography of music itself is generally covered adequately on a national basis.

A few broad conclusions may be drawn about the current status of humanistic bibliography. Due to the proliferation of publication in all fields, bibliographers have had to expand their activities greatly, often with financial assistance from governments and foundations on a scale unprecedented before the middle of this century. The expansion of current bibliographical coverage to several fields which have not hitherto enjoyed it is an encouraging sign; but there are still areas which do not have current bibliographies relating especially to these subjects. The development of monumental bibliographies covering many fields, notably the *MLA International Bibliography*, is a source of some duplication and overlap, but much of it is justified. Political conditions have influenced scholarship as never before, and, in the case of the two Germanies, have been responsible for much wasted effort in the form of duplication of work. One area of partial duplication is in the form of the critical narrative resumés of current research such as *Lustrum* or *Kratylos*; but these organs serve a useful purpose in pulling out the truly relevant material. The coverage of material other than periodical articles and books, notably essays in homage volumes and dissertations, has improved substantially in the last quarter of a century, and the problem now is to keep such bibliographies up to date or to insure that the material appears in other current organs and is adequately indexed. Some fields still suffer from the tardiness of their current bibliographies, but this is an old complaint that extends to all varieties and genres of humanistic scholarship.

It may properly be asked whether existing bibliographies, both serial and nonserial, in the humanities reflect trends in scholarship or even meet the needs of humanistic scholarship today. The answer is, in general, negative. It may safely be stated that no bibliography has ever been universally satisfactory, and this situation is particularly true in the humanities for reasons already indicated. But it is not too difficult to identify areas where improvement is possible, if only time, energy, are funds are available. Perhaps the greatest need is subject indexing in depth. No index can be too detailed. A classified arrangement of a bibliography or an encyclopaedia with only an index of names is rarely

satisfactory. To identify the varied trends in modern textual criticism, to pull together all the outgrowths of the Chomsky school of linguistics, to identify new techniques of programming computers for indexes and concordances is not possible without the expenditure of needless hours or days of effort to wade through a monolithic work such as the *MLA International Bibliography*. The ideal for many specialists is a narrative critical bibliography of the "Year's Work" genre, currently exemplified in such organs as *Lustrum* and *Kratylos*. Even bibliographies without annotation but confined to narrow fields will reflect trends and methods of research to the specialist who knows the field well, but not to the outsider. The "Checklists" in *Studies in Bibliography* mean relatively little to the scholar who is not already *au jour* with the rapid changes in this field that are taking place today.

Perhaps one of the gravest deficiencies of bibliographies of fields international in scope is the failure to cover all countries and all languages with equal effectiveness. There is still a vast corpus of Soviet literature in all fields before the 1960s which was not identified and recorded by bibliographers in western Europe. A cursory examination of *International Bibliography of the History of Religions* and the *Index to Religious Periodical Literature* and of the *Répertoire bibliographique de la philosophie* reveals serious lacunae in coverage of the literature in oriental languages. Such gaps are partially filled by specialized bibliographies of Islamic studies, Indology, Japanology, and so on; but many of these bibliographies are closed books to the student of the history of philosophy or religion on account of the lack of annotation of work in non-European languages. The linguistic arrogance of the European from the Alexandrians on has forced the Oriental to read or even write in Greek, Latin, French, German, or English in different ages. In the meanwhile, those who read only in the latter three languages, or understand bibliographies only in these languages are committing grave errors of omission. We will only know how grave this situation is when the coverage of each "international" bibliography is examined with careful attention to detail.

Humanistic bibliographers can look back with some pride on their achievements since 1945, but the challenge of the future is grave. Above all, the humanities must seek a larger share of government and foundation funds. As long as they are available, they should be divided equitably among the various fields of human endeavor. More international cooperation, particularly in fields that are not limited to national interest such as literatures and language, will provide the bases for bet-

ter planned bibliographical control and make humanistic projects more attractive to those who give financial aid. A careful watch must be kept on technological developments and every effort made to take advantage of them to cut down on drudgery and to expand bibliographical coverage of all varieties. The bibliographer still needs to remind the nonbibliographical scholar that our work provides the sinews of any discipline. To paraphrase Seneca, Biblioteca sine bibliographia mors est, et librorum virorumque sepultura.

Information Obtainable from Analyses of Scientific Bibliographies

R. T. BOTTLE

WHAT MORE can be obtained from bibliographies than access to the primary literature for which they serve as keys? Because they are ready-made collections of documents about a particular facet of science or technology, they are increasingly being used as source material by historians and sociologists of science. Unquestionably they record the (recent) past; can they be used to predict the future? If the answer is even "possibly," then it would be a very valuable exercise to analyze them.

Analyses of bibliographies would provide an alternative to asking the opinion of experts about the likely course of future events, a sophisticated version of which is the so-called Delphi technique. The experts are, of course, familiar with all aspects of a subject, including its literature, and may well have assessed intuitively trends in the literature which will form part of their judgment of that subject's future.

Literature trends can be quantitatively observed without expert knowledge of a specific subject and possibly provide an industrial company, for example, with predictive material without the need to divulge its interest in the subject to an outside expert. These two approaches exhibit distinct analogies with the so-called fundamental and chartist approach to investment analysis—both of which are considered notoriously unreliable! The major difference between the approaches is the time scale involved; while investment analysis charts can point quite quickly to a change in the underlying factors involved, the publication of research is the culmination of a process where the operative decisions were perhaps made several years previously—indeed a decision to abandon further work on the project may have been made long before even a draft manuscript of an eventual publication had been

R. T. Bottle is with the Department of Librarianship, University of Strathclyde, Glasgow, Scotland.

written. In the case of a scientific research program there is usually sufficient momentum established quite early to continue the publication rate for quite some time after it has been decided to terminate it.

Trends observed in bibliographies right now, therefore, will be the reflection of events which occurred some years earlier. They can often be successfully extrapolated a few years into the future; it may even be postulated that, because of the momentum which a research program acquires, they will continue until such time as a constraint affects them (an obvious analogy with Newton's laws of motion). An example of a constraint is a change in policy by a major publisher, and this will be independent of the research scene. As publishers normally plan four or five years ahead, this period represents approximately the maximum time for which extrapolation procedures are valid.

While the literature abounds with qualitative prophecies of doom[1] that science will become buried under its own literature and that communication between research workers will break down and so forth, it is difficult to find many quantitative predictions made some time ago which can now be tested for validity. In an earlier publication[2] I was rash enough to suggest that it *might* be possible to predict the future size of *Chemical Abstracts* from the equation $\log C = 4.415 + 0.0464n$ where C is the number of columns and n is the number of years since 1960. This prediction, which owes much to an earlier study by Strong and Benfey,[3] was followed quite well up to 1962, but then the observed growth changed to a lower rate and was subject to much wider fluctuations than previously (see figure 1). In 1957 Crane intimated by means of a graph that the number of papers abstracted by *Chemical Abstracts* would rise to about 125,000 by 1965.[4] The number actually published in 1965 was nearly 170,000.[5] Perhaps errors of about one-third are acceptable or the best we can expect from such bibliometric predictions projected beyond two or three years. (If Crane could not predict the future of *Chemical Abstracts* any better than this, what hope is there for lesser mortals?)

As a MEDLARS bibliography on kinetocardiography dating back to the 1964 tapes had been provided for a colleague, the number of references by year of publication were counted and noted in table 1. The leader of the research group working on this project was asked if he felt interest was increasing, decreasing or remaining much the same. As his reply was that interest was increasing, it is probable that failure to observe this trend in the number of publications is a result of the Nixon administration's cutbacks in research funding in 1969-70. There are

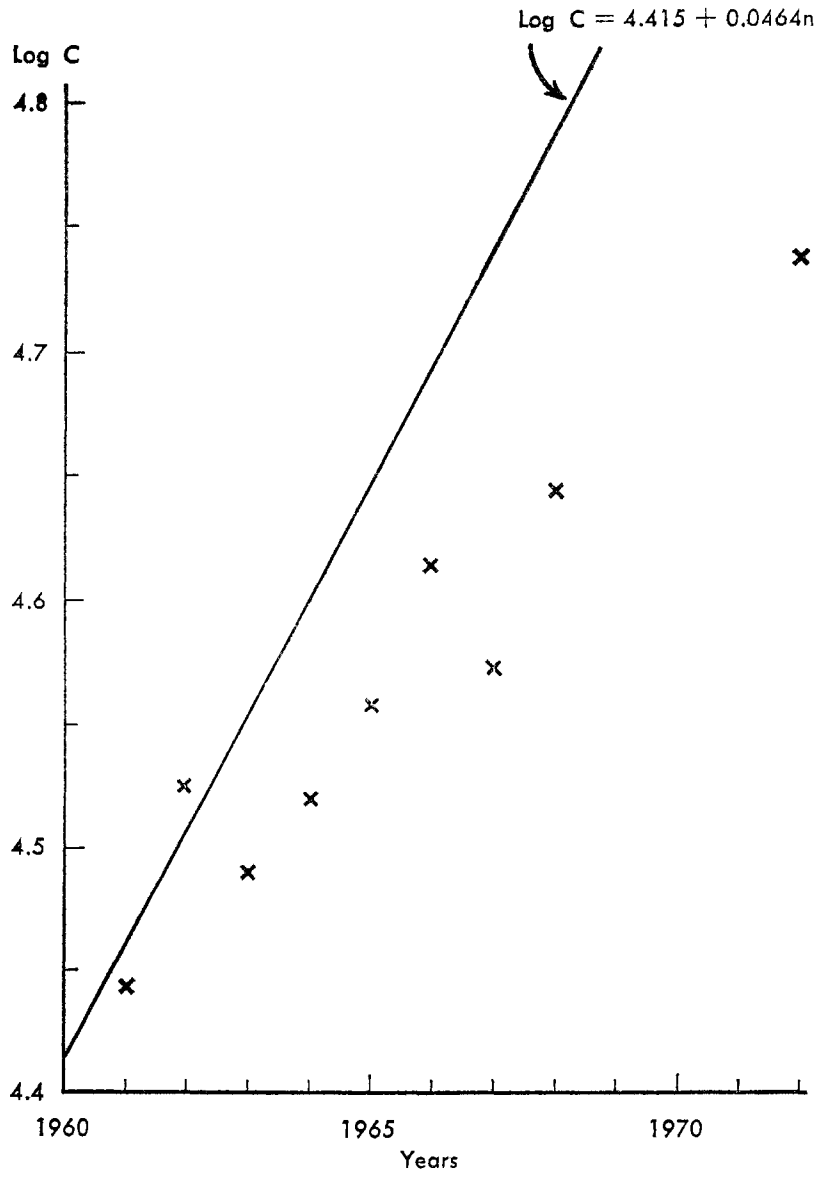

FIG. 1. Predicted and Actual Numbers of Columns in *Chemical Abstracts* vs. Year

then considerable limitations to using this type of study to predict future trends, but there is, however, much other interesting information which can be gained by studying scientific bibliographies, much of which can be of direct practical use to information scientists and their clientele.

A popular myth, especially among scientists, is that de Solla Price's 1963 book *Little Science, Big Science* was the start of interest in studying characteristics of scientific literature. It was undoubtedly a most influential and catalytic book and, in the last few years, there has been

TABLE 1

RESULTS OF MEDLARS SEARCH ON KINETOCARDIOGRAPHY

Year	No. of References
1972	50
1971	66
1970	70
1969	62
1968	72
1967	65
1966	39
1965	39
1964	20

Tapes were searched from 1964 to end of 1972. Because of time lags the data for 1972 (and possibly 1971) are incomplete.

a considerable upsurge in bibliometric studies. It has, however, long been recognized that scientific literature had a structure and properties. Even in 1882 H. C. Bolton proclaimed, "Chemical literature is characterized by two opposing forces, a tendency to dispersal and an effort to collect the widely scattered publications."[5] More recent, perhaps, is the realization that a knowledge of these properties and structure is important for an understanding of information transfer processes in science.

The characteristics of the producers of information and the forms in which it is disseminated are especially important. Bibliographies (and this term is used in this paper to include abstracting and indexing services) can provide the raw data for analyses which seek to quantify these points. The address lists at the back of *Current Contents* are very useful for ascertaining the types of institution in which publishing sci-

entists work and which are the most prolific.[6] A sample of these scientists can then be traced in the author indexes of other bibliographies and their productivity in terms of papers per year can be determined. In most fields academics produce the bulk of the literature, but in computer science Pritchard states that 49.5 percent of authors were employed by firms.[7] A variant of this technique has been used in an attempt to assess the quality of British university chemistry departments using the corporate index to the *Science Citation Index* to identify the productivity of staff members and then determining the number of citations (other than self citations) received by the average staff member.[8]

TABLE 2

LANGUAGE OF PUBLICATION BY SUBJECT

Subject	English	Russian	German	French	Japanese	Czech
Chemistry[9] (*Chemical Abstracts*, 1966)	54.9	21.0	7.1	5.2	3.1	
Physics[10] (*Physics Abstracts*, 1964)	67.7	18.0	6.4	6.1	0.2	
Computer Science[7]	66.5	14.4	7.25	3.175	3.45	1.22
Biology[11] (*Biological Abstracts*, 1967)	68	14	7	5	1	1
Medicine[11] (*Index Medicus*, 1967)	56	3	9	11	3	2
Music & Musicology[12] (*RILM*, 1967)	46	2.5	12.5	9.5	2	4

The above procedures can work quite well in areas where the journal paper accounts for the bulk of the literature. Thus the first stage in analyzing the literature is to determine the proportion which is in the form of papers, reports, patents, theses, etc., and obviously no service which covers only papers can provide this sort of data. For this a bibliography, which has comprehensive coverage of all types of literature, is required. Similarly, a comprehensive bibliography is necessary to study distribution by country of origin or language of publication. Language of publication (especially when matched against the language capabilities of the user) is particularly important to the manager of an information service when planning translation facilities. Details obtained from several bibliometric studies are shown in table 2. Language distributions can change with time; Schwartz and Powers[13] have charted the

decline of German and the rise of English for the primary literature of biochemistry, while in 1970 Webb[14] predicted that, if current trends continued, virtually all biochemical research would be published in English by 1974.

Patents and theses have long been a relatively underused literature form. A detailed investigation of patent literature is currently well advanced at the Polytechnic of North London which will hopefully shed light on patent information/literature relationships. Russian writers have discussed the use of patents statistics in forecasting.[15] A Russian study has shown how "genealogies" of individual engineering ideas can be traced through the reference sections of United States and West German patents.[16] The relationship between theses and the quite considerable quantity of information which is republished from them has recently been discussed.[6]

Bibliographies, if they are reasonably comprehensive, serve to monitor the most readily quantifiable output of science and technology—but, as J. Martyn points out earlier in this issue, a document explosion is not necessarily the same as an information explosion. Nevertheless, bibliographies serve as "catalogs of science and of scientists" and as such can serve as a convenient form of raw data for sociologists of science, information scientists and others. The following examples are not intended as a complete catalog of such studies but merely to illustrate their diversity.

A number of sociological studies on the literature of mathematics, social sciences, etc., are reviewed in Crane's recent book on invisible colleges.[17] Although it contains little of direct practical use to the information scientist, it suggests that the reason for exponential growth of the total number of publications and of new authors publishing for the first time in a new area is due to the interaction of members of a research area with other scientists. If such interaction is absent (as, for example, in English literature), a linear growth rate for these parameters would be observed.

About three-fourths of scientific papers have two or more authors and name ordering patterns in such papers, and their implications, have received considerable attention from sociologists. Zuckerman[18] reports that eminent scientists who become Nobel laureates tend not to insist on being first author to the same extent as do less eminent research team leaders, even five or more years prior to their prize. This concern (and occasional acrimony) over the "social symbolism" of the order of precedence of authors would be much reduced if far more jour-

nal editors laid down rules that authors' names should be in a strict alphabetical sequence. (With a name near the beginning of the alphabet, I can perhaps afford to champion this "democratic" method better than most!)

Creativity as a function of the scientist's age has been much studied. Lehman[19] presented evidence that chemists reached maximum creativity in their mid-thirties. From a bibliometric standpoint it is, of course, easy to measure a scientist's productivity in terms of his publications, but is is a very different matter to evaluate their "worth" or "creativeness" bibliometrically. The criterion of average number of citations per year by other research workers propounded by Matheson[8] is not too easy to obtain; it does not recognize the scientist who, like Flory, has produced an idea ahead of his time, and it overates the one whose work is frequently cited by critics questioning its validity. Although Matheson excluded self-citations, he noted that they ranged from 24 to 46 percent (mean = 31 percent) for the sample of British university chemists investigated. Citation counting as a measure of research achievement has, however, been strongly attacked.[20] It has, however, also been used as a measure of the acceptance of new scientific ideas[21] and as evidence for the operation of the so-called Matthew effect whereby a discovery by an already eminent scientist adds disproportionately more to his prestige than it would to a scientist who has not yet made his mark.

Several areas of science have recently been subjected to bibliometric studies. A good review charting the growth of physics literature and information services was produced by a group at Aslib.[10] Pritchard has characterized the literature of computer science,[7] Meadows and O'Connor have studied the journal literature of astronomy and astrophysics[22] and Simon has produced several studies of biological fields.[23] A different type of study, but nonetheless of immediate value to the librarian, is Mann's compilation of 2,000 journals coded according to the number of food papers produced per year.[24] Thornberry[25] showed that from 1900 to 1956 there was a steady (linear) increase in the proportion of papers on phytovirology relative to those on plant pathology. No tail-off was observed before the study ended.

Well-indexed bibliographies provide a good basis for assessing the information content of document titles and hence the reliability of existing titles indexes, the potential usefulness of a projected one[26] or the relative frequency of occurrence of synonyms in a subject area.[27] Although titles indexes such as *Internationale Bibliographie der Zeit-*

schriftenliteratur, Applied Science & Technology Index, etc., have been in existence for a long time, they do not seem to have been as favorably regarded or as much used by scientists as by librarians; it was not until the 1950s that a combination of factors—such as intolerable delays in the major abstracting services, the introduction of a title index produced by scientists for scientists (the Chemical Society's *Current Chemical Papers*) and possibly Luhn's advocacy of automated titles indexes—induced scientists to regard title indexes seriously. One consequence of this was that they began to pay more attention to the titles which they gave their papers. By the simple technique of counting the keywords per title in different years Tocatlian[28] recently quantified this interesting trend to more informative titles.

The condensation of information ongoing from a primary publication to the secondary literature leads to transmission losses in the several information transfer channels. These losses are either total, due to noninclusion in a specific secondary service, or partial, due to omissions of items of information from a given document by the abstractor or by the indexer. Some methods of estimating transmission losses from studies on bibliographies are discussed elsewhere.[6]

Citation analysis has long been used by librarians with greater or lesser effectiveness for various purposes connected with the management of journal collections. It is an especially useful aid in journal selection policies. A poor example is an analysis of citations in a few American library journals proposed as a basis for journal selection in a library school library.[29] Much more comprehensive data on which to make such a selection were obtained from an analysis of six abstracting services on information science. Although 990 relevant journals were noted, three-fourths of the total information came from just 100 journals.[30] Citation analysis in science has been made easier by the appearance of the *Science Citation Index* and its machine-readable tapes. For example, Martyn and Gilchrist[31] have shown that nearly one-fourth of the citations to British journals were to *Nature* and *The Lancet.* A method of selecting the most prolific journals for a particular topic and for assessing the productivity of journals of marginal interest (provided that they have been indexed by the *Science Citation Index*) has been proposed.[32] These are aspects of the Bradford scattering effect (discussed earlier in this issue) which Douglas[33] and others have suggested is a time-dependent phenomenon and that variations in rankings with time must be taken into account when forecasting a library's future coverage of a field.

Once an abstracting service has got its bibliographic data into machine-readable form, useful statistics can be obtained very readily. (An even greater volume of useless statistics can be churned out just as quickly.) Some of these statistics are helpful in the management of the service,[34] e.g., a ranked list of journals showing their productivity is useful in identifying the core journals of the field. Very detailed summaries of statistical data from *Nuclear Science Abstracts* have been published;[35] one of the many interesting facts produced is that there is a disproportionately high contribution in the area of theoretical physics from the U.S.S.R. BIOSIS often produces analyses of the literature it processes; its annual listings of the percentage of biological serials derived from specific countries is normally to be found in the prefatory pages of the December issues of *Biological Abstracts*. In a study of 1971 literature on nonhuman primates,[36] it was found that of 3,205 papers from 742 primary research publications, more than 45 percent dealt with members of the family Cercopithecidae, but, surprisingly, one-third of the papers made no reference to the specific animal under study. A scattergram by subject concepts and test organisms revealed that the most frequently studied research topic in which nonhuman primates are used is physiology of the nervous system. This has no connection with a previously unpublished study on primates.[37]

Unlike the Russian and French abstracting services, English-language ones are fragmented with respect to the whole field of science and technology. Because of this, considerable overlap occurs. Three of the major services produced in the United States, BioSciences Information Service of Biological Abstracts, Chemical Abstracts Service and *Engineering Index* have recently undertaken a study of the overlap among their respective services. The total number of different journals monitored by the three services in 1970 was 14,592, of which 10,511 were monitored by only one service and only 140 were covered by all three. The biggest overlap in terms of abstracts was between *Engineering Index* and *Chemical Abstracts*—45.3 percent. The overlap of *Chemical Abstracts* with BIOSIS was 28.6 percent.[38] These studies, which are continuing as a "preliminary step in planning for the future" may well have far-reaching effects on users of secondary literature in these areas and on the librarians who provide for their needs.

Errors in bibliographies can occur, and, if not checked with the original, may be perpetuated by subsequent copiers. (Readers familiar with the *Journal of Chemical Education*'s "Textbook Errors" feature will have seen numerous examples where an error of fact was con-

tained in at least two textbooks.) Dobell[39] gives an account of how a title in Czech became substituted for the author's name in *Centralblatt für Bacteriologie und Parasitenkunde* (1887, 1, 537) and for fifty years books and papers on dysentery referred to a paper in an obscure Czech journal by O. Uplavici (Czech for "on dysentery") instead of by the true author, Jaroslav Hlava.[40] There is an apocryphal story of the employee of *Chemisches Zentralblatt,* who, being under notice, managed to insert a nonsense abstract which he attributed to S. C. H. Windler.

In the better bibliographies errors are remarkably few and far between. While I have made no systematic study of this aspect, I did observe only one error in a random sample of 183 subject index entries from the 1967 *Chemical Abstracts.* On the other hand, using the *Science Citation Index* one gets the impression of a rather higher proportion of errors. It is to be expected that as abstracting and indexing services become increasingly produced by automatic means, especially where economic pressures dictate the minimizing of intellectual effort at the input stage, such errors will cause a little, but increasing, frustration for the literature searcher. For example, in a citation index it is clearly impracticable to check every citation in every document processed—one must assume that the journal editor has spotted any errors and inconsistencies; he or she in turn will doubtless assume that authors (or referees) have got the bibliographic details of the citation, such as citer authors' names and initials etc., correct. The *Science Citation Index,* did, however, enter correctly all fifteen citations to Greenstein's joke paper on "armpitin"[41] despite such give away citations as Yolk, A. and White, B., On slicing a hard boiled egg, *Popular Mechanics,* 39:251, 1948 or Goose, M., A child's guide to erotica, Golden Days, Garden City, 1963, etc. The "paper" was also indexed in *Index Medicus* (1966) under "Contraceptives" but I could not find it in either *Chemical Abstracts* or *Biological Abstracts.* A charitable explanation is that perhaps indexers have a sense of humour which abstractors lack.

References

1. An early warning was by Joseph Henry in his Annual Report as Secretary of the Smithsonian Institute in 1851 where he prophesied that scientific knowledge "will begin to totter under its own weight" unless "well digested indexes of subjects not merely referring to volumes or books, but to memoirs, papers and parts of scientific transactions" were produced.

2. Bottle, R. T., ed. *Use of the Chemical Literature.* London, Butterworths, 1962, p. 66.

3. Strong, Laurence E., and Benfey, O. Theodor. "Is Chemical Information Growing Exponentially?" *Journal of Chemical Education*, 37:29-30, Jan. 1960.

4. Crane, E. J. "Chemical Abstracts, Yesterday, Today and Tomorrow," *Proceedings of the Chemical Society*, 1957, pp. 334-40.

5. Bolton, H. Carrington. An address to the American Association for the Advancement of Science, 1882, quoted by Van Patten, Nathan. *Journal of Chemical Education*, 27:431, Aug. 1950.

6. Bottle, R. T. "Scientists, Information Transfer and Literature Characteristics," *Information Scientist*. (In press)

7. Pritchard, Alan. *A Guide to Computer Literature*. 2d ed. London, Bingley, 1972, p. 13.

8. Matheson, A. J. "Centres of Chemical Excellence?" *Chemistry in Britain*, 8:207-10, May 1972.

9. *CAS Today: 60th Anniversary Edition*. Columbus, Ohio, Chemical Abstracts Service, 1967.

10. Anthony, L. J., *et al.* "The Growth of the Literature of Physics," *Reports on Progress in Physics*, 32:709-67, 1969.

11. Wood, D. N. "Foreign Serials and Translations." *In* R. T. Bottle, and H. V. Wyatt, eds. *The Use of Biological Literature*. 2d ed. London, Butterworths, 1971, p. 35.

12. Bottle, R. T., and Chase, W. W. "Some Characteristics of the Literature on Music and Musicology," *Special Libraries*, 63:469-76, Nov. 1972.

13. Schwartz, Ellen S. E., and Powers, Wendell H. "Survey of the Quantity and Distribution of Biochemical Literature," *Journal of Chemical Documentation*, 3:37-42, Jan. 1963.

14. Webb, E. C. "Communication in Biochemistry," *Nature*, 225:132-35, Jan. 10, 1970.

15. Polyakov, G. K. "Patent Statistics: Specifics and Properties," *Voprosy Izobretatel'stva*, 1970:55-60, (1-2) (*Information Science Abstracts*, 6:240, 71-2103, Aug. 1971); and Dauranov, R. B. "Quantitative Studies on the Basis of Patent Materials," *Voprosy Izobretatel'stva*, 1970:49-54, (1-2) (*Information Science Abstracts*, 6:240, 71-2101, Aug. 1971).

16. Korrennoi, A. A. "Quantitative Analysis of Information Relations Between Patents," *Voprosy Izobretatel'stva*, 1969:34-39, (1-2) (*Information Science Abstracts*, 6:178, 71-1601, June 1971).

17. Crane, D. *Invisible Colleges: Diffusion of Knowledge in Scientific Communities*. Chicago, University of Chicago Press, 1972.

18. Zuckerman, Harriet A. "Patterns of Name Ordering Among Authors of Scientific Papers: A Study of Social Symbolism and its Ambiguity," *American Journal of Sociology*, 74:276-91, Nov. 1968.

19. Lehman, Harvey C. "The Chemist's Most Creative Years," *Science*, 127: 1213-22, May 23, 1958.

20. May, Kenneth O. "Abuses of Citation Indexing," *Science*, 156:890-92, May 1967; and Comfort, Alex. "Pop Charts for Science," *Nature*, 227:1069, Sept. 5, 1970.

21. Cole, Stephen. "Professional Standing and the Reception of Scientific Discoveries," *American Journal of Sociology*, 76:286-306, Sept. 1970.

22. Meadows, A. J., and O'Connor, J. G. *A Survey in Depth of a Selected Information Field (Astronomy)* (Final Report, OSTI Report 5092). 1971.

23. Three examples are: Simon, H. R. "Analyses of Bibliographies on Biocontrol," *Journal of Documentation,* 26:337-39, Dec. 1970; ———. "Historische Entwicklung und aktueller Stand der entomologischen Literatur," *Zeitschrift für Pflanzenkrankenheiten und Pflanzenschutz,* 79:413-29, 1972; and ———. "Zur Analyse biologischer Bibliographien," *Heidelberger Jahrbüch,* 15:111-33, 1971.

24. Mann, E. J. *Evaluation of the World Food Literature.* Farnham Royal, Bucks., Commonwealth Agricultural Bureaux, 1967.

25. Thornberry, H. H. "Publications Relative to Vital Diseases in Plants from 1900-1956," *Plant Disease Reporter,* 43:371-73, 1959.

26. Bottle, Robert T., and Seely, Catherine R. "Information Transfer Limitations of Titles of Chemical Documents," *Journal of Chemical Documentation,* 10:256-59, Nov. 1970; and Bottle, R. T. "An Experimental Examination of the Effectiveness of Possible Titles Indexes for Design Literature." Paper No. 9 in *Proceedings of the International Symposium on Information Systems for Designers.* University of Southampton, 1971.

27. Bottle, R. T. "Synonyms, Related Terms and Thesaurus Controlled Indexing." Paper presented at Informatics I, Durham, April 11-13, 1973. (In press)

28. Tocatlian, Jacques Jean. "Are Titles of Chemical Papers Becoming More Informative?" *Journal of the American Society for Information Science,* 21:345-50, Sept. 1970.

29. Reference omitted to avoid embarrassing its author.

30. Novakova, Hana. "The Analysis of World Abstracting Services in Informatics," *Bibliograficky Časopis,* 3:118-36, 1969 (*Information Science Abstracts,* 6:67, 71-550, April 1971).

31. Martyn, John, and Gilchrist, Alan. *An Evaluation of British Scientific Journals* (Occasional Publication, no. 1). London, Aslib, 1968.

32. Bottle, R. T. *The Use of Chemical Literature.* 2d ed. London, Butterworths, 1969, pp. 71-72.

33. Douglas, I. A. "Science Libraries," *Nature,* 239:477, Oct. 20, 1972.

34. Pflueger, Margaret L. "Management Information Obtained from Bibliographic Data," *Proceedings of the American Society for Information Science,* 8:355-61, 1971.

35. Vaden, W. M. An Analysis of World-Wide Contributions to *Nuclear Science Abstracts Vol. 22 (1968)* (TID 25004 and subsequent reports).

36. Kennedy, H. E. "Nonhuman Primate Research Literature—A Subset Analyzed," *Biological Abstracts,* 54(6):xxiv, Sept. 15, 1972.

37. Cantuar, M. "Nonhuman Primates: Attitudes to the Pill," *The Tablet,* 10: 1010, 1969. (Inclusion of this reference in a citation index is inadvisable as it could be regarded as an overspill from Greenstein's paper.)

38. Kennedy, H. E. "Progress Report: BIOSIS/CAS/Ei Overlap Study," *Biological Abstracts,* 53(9):xxiv, May 1, 1972.

39. Dobell, C. "Dr. O. Uplavici (1887-1938)," *Parasitology,* 30:239-41, 1938.

40. Hlava, J. "O Úplavici. Předléžne Sdělení" *Časopis Lékařův Českých,* 26(5), 1887.

41. Greenstein, J. S. "Studies on a New, Peerless Contraceptive Agent: A Preliminary Final Report," *Canadian Medical Association Journal,* 93:1351-55, 1965.

Outlook: The Analyses of Bibliographies in the Future

H. R. SIMON

THE dissemination of knowledge is linked with the general, often exponential growth of products and services, e.g. the Gross National Product. In the realm of knowledge this growth is subject to certain regulations which can also be defined mathematically in an unambiguous way. The more narrowly the area of research is delimited, the more accurately statements can be formulated and trends recognized.

As data for such research, special bibliographies serve with the highest possible degree of completeness;[1] or one may use such collections of literature which point out the most symptomatic characteristics, though they may be incomplete.[2] The latter case represents, in a statistical sense, a random sample.

Trends in special areas of research can be satisfactorily demonstrated in this way. What should be undertaken in a wider context, and particularly for the information sciences, would likewise be an analysis of tendencies with the greatest possible number of quantitative values. In this way it should be possible to answer questions such as:

How many bibliographies will there be in the year 1980, 1990, or 2000?

In what form will they be available (print, microfilm, or computer-stored)?

How are these forms related to each other?

In what ratio to bibliographies are periodicals and reports found, i.e., to how many periodicals and reports does a given bibliography apply?

How does bibliographic control develop, and how great is the overlap in the case of varying degrees of control?

Can the language barrier (particularly as it applies to China and the U.S.S.R.) be minimized or eliminated entirely by effective bibliographic control?

How can these relations be presented quantitatively?

These few examples of questions to which the analyses of bibliographies might hold some answers could be expanded to a catalog of perhaps one hundred questions if one wanted to delineate the entire area of concern. These examples also show clearly that results can only be expected if simulation models are employed. For that reason the computer is indispensable in analyzing bibliographies. Aided by these techniques and the various bibliographies as source material, a collection of prognoses for information science analogous to the models of Meadows and Forrester (M.I.T., Cambridge, Mass.) would have to be drawn up. The various entry parameters and other relational values (e.g., the number of Ph.D.s, the number of periodicals and reports, new periodicals, periodicals which cease publication, redundant publication, institutional grants for the advancement of science) lead us to expect a different configuration for each country. These models would certainly be valuable aids in determining priorities for research and development. In this decision-making process with regard to research and development in the information sciences, purely technical tendencies (their realization and their cost),[3] are at least as important as the development of a general theory for understanding the dissemination of knowledge by these various media. Kochen emphasizes: "The continued survival of the biosphere may well depend on how well individuals in various species can learn to take timely actions on the basis of its collection and growing wisdom."[4]

That this interjection of one's own knowledge is still neither recognized universally nor followed up satisfactorily in research projects is evidenced in the analysis of bibliographies and the determination of their role of analysis in the flow of information in the sciences and the humanities.

In the case of the sciences, methodologically and theoretically comprehensive, far-reaching conceptions are available (cf., contributions by Martyn, Brookes, Bottle and Simon), while in the humanities there is not even a survey of the bibliographies available as sources for analyses of various kinds (cf. especially pp. 58-59 of Thompson in this issue). The difficulties here are unmistakable. Thus a history of science developed from bibliographies could be offered under the title "History of Science Written by Bibliographers." But without use of the computer probably no useful results would be gained. Moreover, the researcher would have to be well-versed in the classical languages and at the same time know as much as possible about computer science. Should these

conditions be fulfilled, a view of history with great objectivity would become possible with such a method, as is demonstrated by de Solla Price.[5]

Besides these briefly sketched new applications of bibliographical analyses, the greatest significance of such analyses may lie in the fact that it will be possible, with these aids, to gain insights into the growth of knowledge and its regularities. Such insights are, for example, now available with the duplication factor of scientific periodicals.[6] Gaining these insights is an even more urgent task since it is postulated today that, on a worldwide basis, science is running on a kind of treadmill.[7] Recognizing this, eliminating it if possible, and thereby saving the entire worldwide political economic structure from ultimate losses remains a nearly hidden objective of bibliographies. Only by means of assured values can a flow of information be interpreted and its tendencies recognized. The length of time prognoses will be of use to us will vary.[8] As stimulus to thought, however, they have an inherent value of their own and should lead to possible revisions in the supply of information.[9]

References

1. Goffman, William. "Mathematical Approach to the Spread of Scientific Ideas —The History of Mast Cell Research," *Nature,* 212:449-52, Oct. 29, 1966.

2. Rothman, Harry, and Woodhead, Michael. "The Use of Citation Counting to Identify Research Trends," *Journal of Documentation,* 27:287-94, Dec. 1971.

3. Weil, B. H. "Document Access," *Journal of Chemical Documentation,* 11:178-85, 1971.

4. Kochen, Manfred. "Stability in the Growth of Knowledge," *American Documentation,* 20:196, July 1969.

5. de Solla Price, Derek J. "Networks of Scientific Papers," *Science,* 149:510-15, July 30, 1965.

6. The literature of synthetic chemistry is growing at a rate of 8.7% a year—that is, doubling every 8.3 years. Garfield, Eugene, *et al.* "The Synthetic Chemical Literature from 1960 to 1969," *Nature,* 242:307-09, March 30, 1973.

7. Arnold, M. "Wachstum und Wissenschaft," *Bild der Wissenschaft,* 9:1164-71, Sept. 1972.

8. Conrad, G. Miles. "Growth of Biological Literature and the Future of Biological Abstracts," *Federation Proceedings,* 16:711-15, Sept. 1957.

9. Brown, W. S., *et al.* "The Future of Scientific Journals," *Science,* 158:1153-59, Dec. 1, 1967; and Simon, H. R. "Die Zunkunft der Biologischen Bibliographie." In *Festschrift für C. Nissen.* (In press)

This Page Intentionally Left Blank

Complete List of Library Trends Issues in Print

			Title	Editor	Date
V.	1,	N. 1	Current Trends in College and University Libraries	R. B. Downs	July 1952
	1	2	Current Trends in Special Libraries	H. H. Henkle	Oct. 1952
	1	3	Current Trends in School Libraries	Alice Lohrer	Jan. 1953
	1	4	Current Trends in Public Libraries	Herbert Goldhor	April 1953
V.	2,	N. 1	Current Trends in Libraries of the U.S. Government	Verner W. Clapp Scott Adams	July 1953
	2	2	Current Trends in Cataloging and Classification	Maurice F. Tauber	Oct. 1953
	2	3	Scientific Management in Libraries	Ralph R. Shaw	Jan. 1954
	2	4	Availability of Library Research Materials	Dorothy M. Crosland William P. Kellam	April 1954
V.	3,	N. 1	Current Trends in Personnel Administration	Bernard Van Horne	July 1954
	3	2	Services to Readers	Leslie W. Dunlap	Oct. 1954
	3	3	Library Associations in the United States and the British Commonwealth	David H. Clift	Jan. 1955
	3	4	Current Acquisitions Trends in American Libraries	Robert Vosper	April 1955
V.	4,	N. 1	Current Trends in National Libraries	David C. Mearns	July 1955
	4	2	Special Materials and Services	Andrew H. Horn	Oct. 1955
	4	3	Conservation of Library Materials	Maurice F. Tauber	Jan. 1956
	4	4	State and Provincial Libraries in the United States and Canada	Paxton P. Price	April 1956
V.	5,	N. 1	American Books Abroad	Dan Lacy Charles Bolte Peter S. Jennison	July 1956
	5	2	Mechanization in Libraries	Arnold H. Trotier	Oct. 1956
	5	3	Manuscripts and Archives	R. W. G. Vail	Jan. 1957
	5	4	Rare Book Libraries and Collections	Howard H. Peckham	April 1957
V.	6,	N. 1	Current Trends in Circulation Services	Wayne S. Yenawine	July 1957
	6	2	Research in Librarianship	A.A.L.S. Committee on Research	Oct. 1957
	6	3	Building Library Resources Through Cooperation	Ralph T. Esterquest	Jan. 1958
	6	4	Legal Aspects of Library Administration	John B. Kaiser	April 1958
V.	7,	N. 1	Current Trends in Book Publishing	Frank L. Schick	July 1958
	7	2	Aspects of Library Public Relations	Len Arnold	Oct. 1958
	7	3	Current Trends in Library Administration	Ernest J. Reece	Jan. 1959
	7	4	Current Trends in Bibliography	Roy B. Stokes	April 1959
V.	8,	N. 1	Current Trends in Adult Education	C. Walter Stone	July 1959
	8	2	Current Trends in Newly Developing Countries	Wilfred J. Plumbe	Oct. 1959
	8	3	Photoduplication in Libraries	James E. Skipper	Jan. 1960
	8	4	Music Libraries and Librarianship	Vincent Duckles	April 1960
V.	9,	N. 1	State Aid to Public Libraries	S. Janice Kee	July 1960
	9	2	Current Trends in Theological Libraries	Niels H. Sonne	Oct. 1960
	9	3	Current Trends in Bookmobiles	Harold Goldstein	Jan. 1961
	9	4	Current Trends in Antiquarian Books	Hellmut Lehmann-Haupt	April 1961
V.	10,	N. 1	Future of Library Service: Demographic Aspects and Implications, Part I	Frank L. Schick	July 1961
	10	2	Future of Library Service: Demographic Aspects and Implications, Part II	Frank L. Schick	Oct. 1961
	10	3	Current Trends in U. S. Periodical Publishing	Helen M. Welch Maurice F. Tauber	Jan. 1962
	10	4	Urban University Libraries	Lorena A. Garloch	April 1962
V.	11,	N. 1	Library Boards	J. Archer Eggen	July 1962
	11	2	Bibliotherapy	Ruth M. Tews	Oct. 1962
	11	3	Law Libraries	Bernita J. Davies	Jan. 1963
	11	4	Financial Administration of Libraries	Ralph H. Parker Paxton P. Price	April 1963
V.	12,	N. 1	Public Library Service to Children	Winifred C. Ladley	July 1963
	12	2	Education for Librarianship Abroad in Selected Countries	Harold Lancour J. Clement Harrison	Oct. 1963
	12	3	Current Trends in Reference Services	Margaret Knox Goggin	Jan. 1964
	12	4	European University Libraries: Current Status and Developments	Robert Vosper	April 1964
V.	13,	N. 1	Research Methods in Librarianship	Guy Garrison	July 1964
	13	2	State and Local History in Libraries	Clyde Walton	Oct. 1964
	13	3	Regional Public Library Systems	Hannis S. Smith	Jan. 1965
	13	4	Library Furniture and Furnishings	Frazer G. Poole	April 1965
V.	14,	N. 1	Metropolitan Public Library Problems Around the World	H. C. Campbell	July 1965
	14	2	Junior College Libraries	Charles L. Trinkner	Oct. 1965
	14	3	Library Service to Industry	Katharine G. Harris Eugene B. Jackson	Jan. 1966
	14	4	Current Trends in Branch Libraries	Andrew Geddes	April 1966

Complete List of Library Trends Issues in Print

		Title	*Editor*	*Date*
V. 15, N.	1	Government Publications	Thomas S. Shaw	July 1966
15	2	Collection Development in University Libraries	Jerrold Orne	Oct. 1966
15	3	Bibliography: Current State and Future Trends. Part 1	Robert B. Downs Frances B. Jenkins	Jan. 1967
15	4	Bibliography: Current State and Future Trends. Part 2	Robert B. Downs Frances B. Jenkins	April 1967
V. 16, N.	1	Cooperative and Centralized Cataloging	Esther J. Piercy Robert L. Talmadge	July 1967
16	2	Library Uses of the New Media of Communication	C. Walter Stone	Oct. 1967
16	3	Abstracting Services	Foster E. Mohrhardt	Jan. 1968
16	4	School Library Services and Administration at the School District Level	Sara K. Srygley	April 1968
V. 17, N.	1	Group Services in Public Libraries	Grace T. Stevenson	July 1968
17	2	Young Adult Service in the Public Library	Audrey Biel	Oct. 1968
17	3	Development in National Documentation and Information Services	H. C. Campbell	Jan. 1969
17	4	The Changing Nature of the School Library	Mae Graham	April 1969
V. 18, N.	1	Trends in College Librarianship	H. Vail Deale	July 1969
18	2	University Library Buildings	David C. Weber	Oct. 1969
18	3	Problems of Acquisition for Research Libraries	Rolland E. Stevens	Jan. 1970
18	4	Issues and Problems in Designing a National Program of Library Automation	Henry J. Dubester	April 1970
V. 19, N.	1	Intellectual Freedom	Everett T. Moore	July 1970
19	2	State and Federal Legislation for Libraries	Alex Ladenson	Oct. 1970
19	3	Book Storage	Mary B. Cassata	Jan. 1971
19	4	New Dimensions in Educational Technology for Multi-Media Centers	Philip Lewis	April 1971
V. 20, N.	1	Personnel Development and Continuing Education in Libraries	Elisabeth W. Stone	July 1971
20	2	Library Programs and Services to the Disadvantaged	Helen H. Lyman	Oct. 1971
20	3	The Influence of American Librarianship Abroad	Cecil K. Byrd	Jan. 1972
20	4	Current Trends in Urban Main Libraries	Larry Earl Bone	April 1972
V. 21, N.	1	Trends in Archival and Reference Collections of Recorded Sound	Gordon Stevenson	July 1972
21	2	Standards for Libraries	Felix E. Hirsch	Oct. 1972
21	3	Library Services to the Aging	Eleanor Phinney	Jan. 1973
21	4	Systems Design and Analysis for Libraries	F. Wilfrid Lancaster	April 1973

Library Trends

Forthcoming numbers are as follows:

October, 1973, *Research in the Fields of Reading and Communication.* Editor: Alice Lohrer, Professor, Graduate School of Library Science, University of Illinois, Urbana-Champaign.

January, 1974, *Evaluation of Library Services.* Editor: Sarah Reed, Associate Dean, Graduate Library School, Indiana University, Bloomington, Indiana.

April, 1974, *Science Materials for Children and Young People.* Editor: George S. Bonn, Professor, Graduate School of Library Science, University of Illinois, Urbana-Champaign.

July, 1974, *Health Sciences Libraries.* Editor: Joan Titley, Health Sciences Librarian, University of Louisville, Louisville, Kentucky.